

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
24 December 2003 (24.12.2003)

PCT

(10) International Publication Number  
**WO 03/107009 A2**

(51) International Patent Classification<sup>7</sup>: **G01N 33/68**,  
33/53

Alto, CA 94303 (US). **MURRAY, Christopher, J.**  
[US/US]; 221 Coldbrook Lane, Soquel, CA 95073 (US).

(21) International Application Number: PCT/US03/18187

(74) Agent: **ANDERTON, H., Thomas, Jr.**; Genencor International, Inc., 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).

(22) International Filing Date: 9 June 2003 (09.06.2003)

(25) Filing Language: English

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(26) Publication Language: English

(30) Priority Data:  
60/388,386 12 June 2002 (12.06.2002) US

(71) Applicant (*for all designated States except US*): **GENENCOR INTERNATIONAL, INC.** [US/US]; 925 Page Mill Road, Palo Alto, CA 94304 (US).

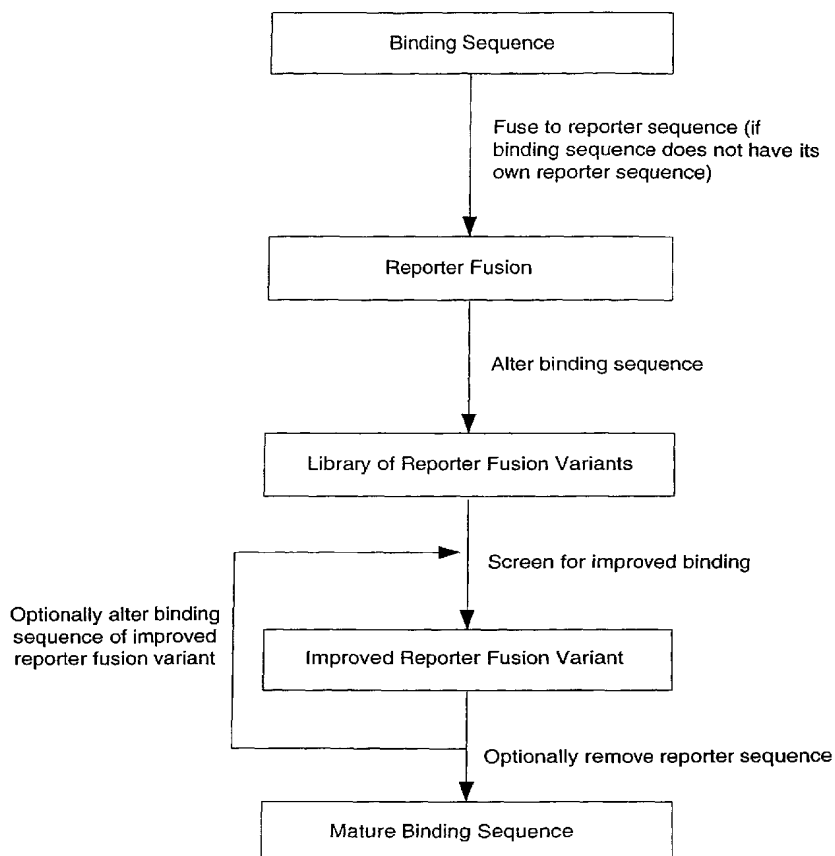
(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): **SCHELLENBERGER, Volker** [US/US]; 914 Moreno Avenue, Palo

[Continued on next page]

(54) Title: METHODS FOR IMPROVING A BINDING CHARACTERISTIC OF A MOLECULE



(57) Abstract: The present invention relates to methods for improving a binding characteristic of a molecule, *e.g.*, a peptide, for a binding target, in which the molecule is covalently linked to a detectable moiety, *e.g.*, an enzyme, or an active portion or derivative thereof. The present invention also relates to molecules produced by the methods of the present invention.



WO 03/107009 A2



SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

— *without international search report and to be republished upon receipt of that report*

## METHODS FOR IMPROVING A BINDING CHARACTERISTIC OF A MOLECULE

---

### 1. FIELD OF THE INVENTION

The present invention relates to methods for improving a binding characteristic of a molecule, *e.g.*, a peptide, for a binding target, in which the molecule is joined to a detectable moiety, *e.g.*, an enzyme, or the active portion thereof. The present invention also relates to molecules produced by the methods of the present invention.

### 2. BACKGROUND OF THE INVENTION

Molecules that bind a particular target are useful in a number of applications, including diagnostic and therapeutic methods, affinity purification methods, methods of delivery to specific locations, *etc.* Of particular interest are proteins and peptides, including antibodies, that bind particular and specific targets, for example, nucleic acids or proteins. Molecules that have particular binding abilities can be generated in a number of ways. One method is by immunizing an animal with a target molecule and subsequently isolating antibodies, or fragments thereof, that bind the target. Another method is by using phage display or other display methods to isolate binding molecules, including proteins. *See, e.g.*, Chen *et al.*, 2001, *Nat. Biotechnol.* 19:537-42.

In many cases, the binding properties of the isolated molecules obtained by such methods are not ideal for their ultimate application. Several methods have been described to improve the binding properties, which mostly involve generating variants of the starting sequence and identifying variants with improved binding properties. *See, e.g.*, Yang *et al.*, 1995, *J. Mol. Biol.* 254:392-403; Schier *et al.*, 1996, *J. Mol. Biol.* 263:551-67; and Beiboer *et al.*, 2000, *J. Mol. Biol.* 296:833-49, which are related to phage display-based methods. However, such methods are time consuming and require several rounds of “panning”. Further, it is known that such methods can result in the enrichment of binding molecules that show reduced binding affinity for the selected target. Thus, typically, tens of thousands of potential molecules must be screened to isolate those with improved binding ability, and this screening process typically requires the use of helper reagents, such as anti-phage antibodies and antibody-enzyme conjugates, that limit the sensitivity and precision of subsequent screens.

Another frequently used method for screening is the ELISA method. In this method, a binding target is attached to a surface (*e.g.*, a well in a microtiter dish). The target attached to the surface is incubated with a candidate binding molecule in a first binding

reaction. The first binding reaction is washed to remove unbound candidate molecules. A helper reagent (*e.g.*, an antibody-enzyme conjugate) is then added for a second binding reaction. The helper reagent binds the candidate molecules bound to the target. After a second wash to remove unbound helper reagent, a substrate is added. The substrate is converted into a detectable form by helper reagent bound to candidate binding molecules bound to target.

The ELISA methodology has several drawbacks, principally due to the requirement of a second binding reaction. During the second binding reaction, the helper reagent can interact non-specifically with the target or reaction vessel thus leading to a high background signal, which limits the ability to detect weakly bound molecules.

Other assay formats are also used to measure or detect binding interactions, including radioimmune and biotinylation-based binding assays. Similarly, these assays also require helper reagents and suffer from the same limitations.

Thus, there remains in the art a need for more sensitive and efficient methods for identifying molecules with improved binding characteristics.

Citation of a reference in this or any section of the specification shall not be construed as an admission that such reference is prior art to the present invention.

### 3. SUMMARY OF THE INVENTION

The present invention relates to compositions and methods for improving a binding characteristic of a binding molecule, *e.g.*, a peptide binding sequence, in which the binding molecule is joined to a reporter molecule, *e.g.*, an enzyme, or the active portion or catalytic domain thereof. The reporter molecule, and thus a binding molecule linked to it, can be detected without a second binding reaction, *e.g.*, of the type used in standard ELISA assays, as illustrated in Figure 4.

In one aspect, the present invention provides methods of improving a binding characteristic, *e.g.*, affinity, selectivity, release rate or turnover rate, of a prototype binding molecule for a target, comprising: contacting the target with a reporter fusion under conditions that allow the reporter fusion to bind to the target, wherein the reporter fusion comprises a reporter molecule and a variant binding molecule derived from the prototype binding molecule that binds to the target, and selecting the reporter fusion if it binds the target with an improved binding characteristic relative to that of the prototype binding domain for the target.

In another aspect, the invention provides methods for improving a binding characteristic of a binding molecule for a target, comprising: contacting the target with a

library comprising a multiplicity of reporter fusions, under conditions that allow a reporter fusion to bind the target, wherein said reporter fusions comprise a reporter molecule and a variant binding molecule derived from a prototype binding molecule that binds the target, and selecting a reporter fusion that binds the target and has a binding characteristic for the target that is improved relative to the binding characteristic of the prototype binding molecule for the target.

In another aspect, the invention provides methods for improving a binding characteristic of a binding molecule for a target, comprising: contacting a target with a library comprising a multiplicity of reporter fusions, under conditions that allow a reporter fusion to bind the target, wherein said reporter fusions comprise a reporter molecule and a variant binding molecule derived from a prototype binding molecule that binds the target, selecting a reporter fusion bound to the target and having a binding characteristic that is improved relative to the binding characteristic of the prototype binding molecule, and removing the reporter molecule from the selected reporter fusion.

In one embodiment, the selecting step comprises incubating the reporter fusion bound to the target under conditions that cause a reporter fusion with an undesirable binding characteristic to dissociate from the target. In another embodiment, the selecting step comprises incubating the reporter fusion bound to the target with multiple rounds of conditions that cause a reporter fusion with an undesirable binding characteristic to dissociate from the target, wherein each subsequent round of conditions causes dissociation of a reporter fusion with a better binding characteristic for the target than the previous round. In another embodiment, the amount of bound reporter fusion is measured between one or more rounds of dissociation. In another embodiment, the selecting step comprises selecting a reporter fusion if it binds to the target under a first condition better than it binds to the target under a second condition. In another embodiment, the condition is pH, with the first condition being a pH lower than the second condition. In another such embodiment, the first condition is a pH higher than the second condition. In another embodiment, the condition is temperature, with the first condition being a lower temperature than the second condition. In another embodiment, the first condition is a temperature higher than the second condition. In another embodiment, the selecting step comprises incubating the reporter fusion bound to the target in the presence of proteases or under conditions that degrade or destabilize the reporter fusion. In another embodiment, conditions may include, but are not limited to, heat, pH or subjugation to solutes that affect stability.

In another embodiment, the method further comprises repeating the contacting and selecting steps, wherein the variant binding molecule selected in a previous selection step is the prototype binding molecule of a subsequent contacting step.

In another embodiment, the reporter molecule is a reporter sequence. In another embodiment, the reporter sequence is an enzyme or a functional fragment or derivative of an enzyme. In another embodiment, the enzyme is a  $\beta$ -lactamase,  $\beta$ -galactosidase, phosphatase, peroxidase, reductase, esterase, hydrolase, isomerase or protease.

In another embodiment, the reporter fusion is selected if it has a binding affinity that is greater than the binding affinity of the prototype binding molecule. In another embodiment, the reporter fusion is selected if it has a binding affinity that is less than the binding affinity for the target of the prototype binding molecule. In another embodiment, the reporter fusion is selected if it has a binding selectivity that is greater than the binding selectivity of the prototype binding molecule. In another embodiment, the reporter fusion is selected if it has a binding selectivity that is less than the binding selectivity of the prototype binding molecule. In another embodiment, the reporter fusion is selected if it has a release rate that is greater than the release rate of the prototype binding molecule. In another embodiment, the reporter fusion is selected if it has a release rate that is less than the release rate of the prototype binding molecule. In another embodiment, the reporter fusion is selected if it has a turnover rate that is greater than the turnover rate of the prototype binding molecule. In another embodiment, the reporter fusion is selected if it has a turnover rate that is less than the turnover rate of the prototype binding molecule.

In another embodiment, the prototype binding molecule binds the target with a  $K_d$  of about 100  $\mu$ M or less, 10  $\mu$ M or less, 1  $\mu$ M or less, 100 nM or less, about 90 nM or less, about 80 nM or less, about 70 nM or less, about 60 nM or less, about 50 nM or less, about 40 nM or less, about 30 nM or less, about 20 nM or less, about 10 nM or less, about 5 nM or less, about 1 nM or less or about 0.1 nM or less. In yet another embodiment, the selected variant sequence binds the target with a  $K_d$  of about 100  $\mu$ M or less, 10  $\mu$ M or less, 1  $\mu$ M or less, 100 nM or less, about 90 nM or less, about 80 nM or less, about 70 nM or less, about 60 nM or less, about 50 nM or less, about 40 nM or less, about 30 nM or less, about 20 nM or less, about 10 nM or less, about 5 nM or less, about 1 nM or less or about 0.1 nM or less. In another embodiment, the selected variant sequence binds the target with a  $K_d$  of about 100  $\mu$ M or more, 10  $\mu$ M or more, 1  $\mu$ M or more, 100 nM or more, about 90 nM or more, about 80 nM or more, about 70 nM or more, about 60 nM or more, about 50 nM or more, about 40 nM or

more, about 30 nM or more, about 20 nM or more, about 10 nM or more, about 5 nM or more, about 1 nM or more or about 0.1 nM or more.

In another embodiment, the variant binding molecule has been covalently modified relative to the prototype binding molecule.

5 In another embodiment, the binding molecule is a binding sequence. In another embodiment, the variant binding sequence has an amino acid sequence that is at least 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 98 or 99% identical to the amino acid sequence of the prototype binding sequence. In another embodiment, the variant binding sequence has been post-translationally modified relative to the prototype binding sequence.

10 In another aspect, the present invention provides a binding molecule produced or identified by the methods of the present invention.

The present invention can be more fully explained by reference to the following drawings, detailed description and illustrative examples.

#### 15 4. BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates an embodiment of the present invention.

Figure 2 illustrates a reporter fusion and a target of the invention.

Figure 3 illustrates an embodiment of the invention by which variant binding sequences are selected on the basis of having a particular dissociation constant that  
20 corresponds to tighter binding and a slow dissociation of the binding sequence and the target.

Figure 4 illustrates certain differences between a method of the present invention and a standard ELISA assay.

Figure 5 illustrates an embodiment of the invention for selecting pH-dependent binding sequences.

25 Figure 6 illustrates an embodiment of the invention for screening reporter fusions on a surface.

Figure 7 illustrates an amino acid sequence of  $\beta$ -lactamase.

Figure 8 presents a schematic diagram of plasmid pADEPT06. P lac = lac promoter, Pel B leader sequence = signal seq, L49VH = Heavy chain, L49VL = Light chain,  
30 218 linker = linker region between heavy and light chains,  $\beta$ -lactamase =  $\beta$ -lactamase gene, L49sFv-bl = scFv-BLA fusion, CAT = chloramphenicol resistance gene.

Figure 9 shows results of a secondary screening of 21 mutants in quadruplicates. The x-axis shows variant designation and the y-axis shows the performance index. A ratio of bound activity at  $T_1$  vs.  $T_0$  was calculated for each mutant, and the

performance index was calculated by dividing the ratio of mutant over parent, as shown in Table 3.

Figure 10 present details related to plasmid pME27.1 Figure 10A presents a schematic diagram of plasmid pME27.1. P lac = lac promoter, Pel B leader sequence = signal seq, CAB1scFv=single chain antibody, BLA=  $\beta$ -lactamase gene, CAT = chloramphenicol resistance gene, T7 terminator=terminator. Figure 10B presents shows the sequence of CAB1-scFv, the CDRs and mutations chosen for combinatorial mutagenesis. Figure 10C presents and nucleotide sequence of pME27.1 Figure 10D shows the amino acid sequence of CAB1 which shows, for example, the sequence of the heavy chain, the sequence of the linker, the sequence of the light chain and the sequence of BLA.

Figure 11 shows binding assays and SDS page results. Specifically, Figure 11A shows the binding of variants from library NA05; Figure 11B displays and SDS PAGE of stable CAB1-BLA variants of the NA05 library; Figure 11C shows binding of various isolates from NA06 to CEA.

Figure 12 shows a comparison of vH and vL sequences of CAB1-scFv with a published frequency analysis of human antibodies. Specifically, Figure 12A shows the observed frequencies of the five most abundant amino acids in alignment of human sequence in the heavy chain; Figure 12B shows the observed frequencies of the five most abundant amino acids in alignment of human sequence in the light chain.

Figure 13 shows screening results of NA08 library. The x-axis shows binding at pH 7.4, and the Y-axis shows binding at pH 6.5. Clones that were chosen are represented by a square.

Figure 14 shows positions that were chosen for combinatorial mutagenesis.

Figure 15 shows pH-dependent binding of NA08 variants to immobilized carcinoembryonic antigen.

Figure 16 shows a chromatogram for 18 hour old extract of ME27.1

Figure 17 shows an SDS-PAGE for 18 hour extract of ME27.1

Figure 18 shows a chromatogram for 26 hour old extract of ME27.1

Figure 19 shows an SDS-PAGE for 26 hour extract of ME27.1

Figure 20 shows a chromatogram for 26 hour extract for ME 27.1 (4-5 days)

Figure 21 shows an SDS-PAGE for 26 hour old extract. Conditions: 4-12% Tris-Bis/MES/Reducing conditions.

Figure 22 shows CAB1 purification using anion exchange and PBA.



## 5. DETAILED DESCRIPTION OF THE INVENTION

A “binding molecule,” unless otherwise stated, includes both “prototype binding molecules” and “variant binding molecules,” for example, a binding sequence.

5 A “prototype binding molecule” is a molecule that has a measurable binding affinity for a target of interest.

A “variant binding molecule” is a molecule that is similar to, but different from, a prototype binding molecule. The difference can be, for example, any difference in structure, including, *e.g.*, the addition, deletion or substitution of one or more atoms, amino acid residues or functional groups.

10 A “binding sequence,” unless otherwise stated, includes both “prototype binding sequences” and “variant binding sequences.”

A “prototype binding sequence” is a peptide, polypeptide or protein sequence that has a measurable binding affinity for a target of interest.

15 A “variant binding sequence” is a peptide, polypeptide or protein sequence that is similar to, but different from, a prototype binding sequence. The difference can be, for example, any difference in sequence, including, *e.g.*, addition, substitution, and/or deletion of one or more amino acids. The difference also can be or include any form of covalent modification, *e.g.*, a post-translational modification.

20 A “target” is anything, or any combination of things, to which a peptide, polypeptide or protein can bind.

A “reporter molecule” is a molecule that can be detected independent of its binding to a detectable molecule, for example, a labeled antibody or other reporter molecule-binding molecule. A reporter sequence is a type of reporter molecule.

25 A “detectable molecule” is a macromolecule that may bind to a molecule and may be used to detect a reporter molecule; a detectable molecule is not a small molecule substrate.

30 A “reporter sequence” is a reporter molecule that comprises a peptide, polypeptide or protein sequence that can be detected independent of its binding to a detectable molecule, for example, a labeled antibody or other reporter-sequence binding molecule. The reporter sequence can be, for example, an enzyme, a catalytically active fragment or derivative of a protein, or a labeled peptide, polypeptide or protein, *e.g.*, a fluorescently labeled or a radioactively labeled peptide, polypeptide or protein.

A “reporter fusion” is a molecule having a binding molecule and a reporter molecule that are bound to each other, *e.g.*, covalently bound to each other. The reporter

fusion can optionally comprise other elements, for example, one or more linker molecules joining one or more parts of the reporter fusion, *e.g.*, a reporter molecule and a binding molecule. Examples of reporter fusions include chimeric polypeptides and targeted enzymes as described in U.S. Pat. App. Ser. No.s 10/022,073 and 10/022,097, both filed December 13, 2001, and incorporated herein by reference in their entireties.

A "binding characteristic" is a measure of the interaction of two molecules. Examples of binding characteristics include affinity, selectivity, release rate, turnover rate, stability of a molecule necessary for binding and purification of a molecule which has additional or other binding characteristics. Turnover may refer to *in vitro* or *in vivo* internalization and/or degradation by a cell or tissue of any or all of the molecules engaged in the binding interaction that renders the molecules unavailable for binding. For example, a target and/or binding molecule may be internalized by a cell and degraded intracellularly. Binding molecules and targets may also be degraded by cell-surface proteases. Alternatively, following internalization, any or all of the binding molecules may be exocytosed to the cell surface and be accessible and available for binding interactions.

"Selectivity" describes the ability of a binding molecule to discriminate between different targets. A binding molecule is said to have high selectivity if it binds with significantly higher affinity to its intended target than to most other surfaces or molecules.

Unless otherwise noted, the term "protein" is used interchangeably here with the terms "peptide" and "polypeptide," and refers to a molecule comprising two or more amino acid residues joined by a peptide bond.

The terms "cell", "cell line", and "cell culture" can be used interchangeably and all such designations include progeny. Thus, the words "transformants" or "transformed cells" include the primary transformed cell and cultures derived from that cell without regard to the number of transfers. All progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Mutant progeny that have the same functionality as screened for in the originally transformed cell are included in the definition of transformants. The cells can be prokaryotic or eukaryotic.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for procaryotes, for example, include a promoter, optionally an operator sequence, a ribosome binding site, positive retroregulatory elements (see, *e.g.*, U.S. Pat. No. 4,666,848, incorporated herein by reference), and possibly other sequences. Eucaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

The term "expression clone" refers to DNA sequences containing a desired coding sequence and control sequences in operable linkage, so that hosts transformed with these sequences are capable of producing the encoded proteins. The term "expression system" refers to a host transformed with an expression clone. To effect transformation, the expression clone may be included on a vector; however, the relevant DNA may also be integrated into the host chromosome.

The term "gene" refers to a DNA sequence that comprises control and coding sequences necessary for the production of a protein, polypeptide or precursor.

The term "operably linked" refers to the positioning of the coding sequence such that control sequences will function to drive expression of the protein encoded by the coding sequence. Thus, a coding sequence "operably linked" to control sequences refers to a configuration wherein the coding sequences can be expressed under the direction of a control sequence.

The term "oligonucleotide" as used herein is defined as a molecule comprised of two or more deoxyribonucleotides or ribonucleotides. The exact size will depend on many factors, which in turn depends on the ultimate function or use of the oligonucleotide. Oligonucleotides can be prepared by any suitable method, including, for example, cloning and restriction of appropriate sequences and direct chemical synthesis by a method such as the phosphotriester method of Narang et al., 1979, Meth. Enzymol. 68:90-99; the phosphodiester method of Brown et al., 1979, Meth. Enzymol. 68:109-151; the diethylphosphoramidite method of Beaucage et al., 1981, Tetrahedron Lett. 22:1859-1862; and the solid support method of U.S. Pat. No. 4,458,066, each incorporated herein by reference. A review of synthesis methods is provided in Goodchild, 1990, Bioconjugate Chemistry 1(3):165-187, incorporated herein by reference.

The term "primer" as used herein refers to an oligonucleotide which is capable of acting as a point of initiation of synthesis when placed under conditions in which primer extension is initiated. Synthesis of a primer extension product that is complementary to a nucleic acid strand is initiated in the presence of the requisite four different nucleoside triphosphates and a DNA polymerase in an appropriate buffer at a suitable temperature. A "buffer" includes cofactors (such as divalent metal ions) and salt (to provide the appropriate ionic strength), adjusted to the desired pH.

A primer that hybridizes to the non-coding strand of a gene sequence (equivalently, is a subsequence of the coding strand) is referred to herein as an "upstream" or

“forward” primer. A primer that hybridizes to the coding strand of a gene sequence is referred to herein as an “downstream” or “reverse” primer.

The terms “restriction endonucleases” and “restriction enzymes” refer to enzymes, typically bacterial in origin, which cut double-stranded DNA at or near a specific nucleotide sequence.

Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, asparagine, glutamine, serine, threonine, tyrosine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan, cysteine, glycine), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Standard three-letter or one-letter amino acid abbreviations are used herein.

The peptides, polypeptides and proteins of the invention can comprise one or more non-classical amino acids. Non-classical amino acids include but are not limited to the D-isomers of the common amino acids,  $\alpha$ -amino isobutyric acid, 4-aminobutyric acid (4-Abu), 2-aminobutyric acid (2- Abu), 6-amino hexanoic acid (Ahx), 2-amino isobutyric acid (2-Aib), 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine,  $\beta$ -alanine, fluoro-amino acids, designer amino acids such as  $\beta$ -methyl amino acids,  $\text{Ca}$ -methyl amino acids,  $\text{Na}$ -methyl amino acids, and amino acid analogs in general.

As used herein, a “point mutation” in an amino acid sequence refers to either a single amino acid substitution, a single amino acid insertion or single amino acid deletion. A point mutation preferably is introduced into an amino acid sequence by a suitable codon change in the encoding DNA. Individual amino acids in a sequence are represented herein as  $\text{A}_N$ , wherein A is the standard one letter symbol for the amino acid in the sequence, and N is the position in the sequence. Mutations within an amino acid sequence are represented herein as  $\text{A}_1 \text{NA}_2$ , wherein  $\text{A}_1$  is the standard one letter symbol for the amino acid in the unmutated protein sequence,  $\text{A}_2$  is the standard one letter symbol for the amino acid in the mutated protein sequence, and N is the position in the amino acid sequence. For example, a G46D mutation represents a change from glycine to aspartic acid at amino acid position 46. The amino acid positions are numbered based on the full-length sequence of the protein from which the region encompassing the mutation is derived. Representations of nucleotides and point mutations in DNA sequences are analogous.

As used herein, a “chimeric” protein refers to a protein whose amino acid sequence represents a fusion product of subsequences of the amino acid sequences from at least two distinct proteins. A chimeric protein preferably is not produced by direct manipulation of amino acid sequences, but, rather, is expressed from a “chimeric” gene that encodes the chimeric amino acid sequence.

The term “host immune response” refers to a response of a host organism’s immune system to contact with an immunogenic substance. Specific aspects of a host immune response can include, *e.g.*, increased antibody production, T cell activation, monocyte activation or granulocyte activation. Each of these aspects can be detected and/or measured using standard *in vivo* or *in vitro* methods.

The term “Ab” or “antibody” refers to polyclonal and monoclonal antibodies, an entire immunoglobulin or antibody or any functional fragment of an immunoglobulin molecule that binds to the target antigen. Examples of such functional entities include complete antibody molecules, antibody fragments, such as Fv, single chain Fv, complementarity determining regions (CDRs), V<sub>L</sub> (light chain variable region), V<sub>H</sub> (heavy chain variable region), and any combination of those or any other functional portion of an immunoglobulin peptide capable of binding to target antigen.

The term “% sequence homology” is used interchangeably herein with the terms “% homology,” “% sequence identity” and “% identity” and refers to the level of amino acid sequence identity between two or more peptide sequences, when aligned using a sequence alignment program. For example, as used herein, 80% homology means the same thing as 80% sequence identity determined by a defined algorithm, and accordingly a homologue of a given sequence has greater than 80% sequence identity over a length of the given sequence. Exemplary levels of sequence identity include, but are not limited to, 60, 70, 80, 85, 90, 95, 98 or 99% or more sequence identity to a given sequence.

Exemplary computer programs which can be used to determine identity between two sequences include, but are not limited to, the suite of BLAST programs, *e.g.*, BLASTN, BLASTX, and TBLASTX, BLASTP and TBLASTN, known to one of skill in the art and publicly available on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. See also Altschul *et al.*, 1990, *J. Mol. Biol.* 215: 403-10 (with special reference to the published default setting, *i.e.*, parameters *w*=4, *t*=17) and Altschul *et al.*, 1997, *Nucleic Acids Res.*, 25:3389-3402. Sequence searches are typically carried out using the BLASTP program when evaluating a given amino acid sequence relative to amino acid sequences in the GenBank Protein Sequences and other public databases. The BLASTX program is preferred for

searching nucleic acid sequences that have been translated in all reading frames against amino acid sequences in the GenBank Protein Sequences and other public databases. Both BLASTP and BLASTX are run using default parameters of an open gap penalty of 11.0, and an extended gap penalty of 1.0, and utilize the BLOSUM-62 matrix. *See Altschul, et al., 1997.*

5                   A preferred alignment of selected sequences in order to determine “% identity” between two or more sequences, is performed using for example, the CLUSTAL-W program in MacVector version 6.5, operated with default parameters, including an open gap penalty of 10.0, an extended gap penalty of 0.1, and a BLOSUM 30 similarity matrix.

                  “Hit density” is the fraction of useful clones in a library.

10                  “Hapaxomer” is a restriction endonuclease that generates unique ends. *See Berger, S. L. Anal Biochem 222:1 (1994).*

                  The present invention relates to methods and compositions for identifying variants of a binding molecule that have improved binding properties. The methods use reporter fusions comprising variant binding molecules and reporter molecules. Variants with  
15 improved binding properties are identified using the reporter molecule. The process can be repeated multiple times. Ultimately, the binding molecule can be produced without its reporter, either alone or as part of a larger molecule, *e.g.*, a binding sequence that is part of a larger polypeptide. One embodiment of the method is illustrated in Figure 1.

### 20                   Reporter Fusions

                  A reporter fusion comprises a binding molecule operably linked to a reporter molecule. The binding molecule and the reporter molecule are operably linked if the binding molecule can bind the target and the reporter molecule can be detected. In one embodiment, the reporter fusion comprises a plurality of binding molecules. In another embodiment, the  
25 reporter fusion comprises a plurality of reporter molecules. In another embodiment, the reporter fusion comprises a plurality of binding molecules and a plurality of reporter molecules.

                  The binding molecule and the reporter molecule can be joined together using any means for doing so provided that the binding molecule is able to bind the target and the  
30 reporter molecule is detectable. In one embodiment, the binding molecule and the reporter molecule are covalently attached, for example, covalently attached to each other directly (*e.g.*, through a peptide bond or a disulfide bond), or covalently attached to each other via a linker. Examples of linkers include peptides and peptide analogs (*e.g.*, peptide nucleic acids), nucleic acids and nucleic acid analogs, and chemical cross linkers such as p-azidobenzoyl hydrazide,

N-(4-(p-azidosalicylamido)butyl)-3-(2'-pyridylthio)-propionamide,  
 1-(p-azidosalicylamido)-4-(iodoacetamido)butane, 4-(p-azidosalicylamido)butylamine,  
 4,4'-diazidodiphenyl-ethane, 4,4'-diazidodiphenyl-ether, dithio bis phenyl azide,  
 bis(b-(4-azidosalicylaminoethyl)disulfide,  
 5 sulfosuccinimidyl-6(4'-azido-2'-nitrophenylamino)hexanoate,  
 sulfosuccinimidyl-4(p-azidophenyl)butyrate,  
 sulfosuccinimidyl-(4-azidosalicylamido)hexanoate, N-hydroxysuccinimidyl-4-azido benzoate,  
 N-hydroxysulfosuccinimidyl-4-azido benzoate,  
 sulfosuccinimidyl-2-(p-azidosalicylamido)-ethyl-1,3-dithiopropionate, p-azidophenyl glyoxal,  
 10 monohydrate, N-(4-(p-azidosalicylamido)butyl)-3-(2'-pyridylthio)-propionamide, and  
 1-(p-azidosalicylamido)-4-(iodoacetamido)butane. In a more particularly defined  
 embodiment, the binding molecule is a binding sequence, the reporter molecule is a reporter  
 sequence, and the reporter fusion is a fusion protein. The fusion protein can be synthesized  
 chemically, by direct manipulation and joining of peptides, or translated *in vivo* or *in vitro*  
 15 from an appropriate nucleic acid template, as described below. Examples of reporter fusions  
 that are fusion proteins are provided in, for example, Yamabhai *et al.*, 1997, Anal. Biochem.  
 247:143-51; Schlehuber *et al.*, 2001, Biophys. Chem. 96:213-28; Griep *et al.*, 1999, Prot.  
 Express. Purif. 16:63-69; Morino *et al.*, 2001, J. Immunol. Meth. 257:175-84; Wright *et al.*,  
 2001, 253:223-32, incorporated herein by reference in their entireties.

20 In one aspect, the present invention provides a library comprising a multiplicity  
 of reporter fusions. Various reporter fusions in the library comprise a reporter sequence and a  
 different variant binding molecule. The variant sequences are similar to, but different from, a  
 prototype binding molecule. In one embodiment, the variant binding molecules are generated  
 from a reporter fusion comprising the reporter molecule and the prototype binding molecule.  
 25 In another embodiment, the reporter fusion comprises a polypeptide comprising a reporter  
 sequence and a binding sequence. In a more particularly defined embodiment, variant binding  
 sequences are generated by mutating a nucleic acid encoding the reporter fusion. The  
 mutagenesis can target all or part of the prototype binding sequence, and can alter the  
 prototype binding sequence in any way, including, for example, adding, deleting or  
 30 substituting one or more amino acids. If more than one change is made, they can be made  
 contiguously or in different parts of the prototype binding sequence.

In another embodiment, the reporter fusion comprises the binding sequence  
 and/or the reporter sequence as an integral component. This approach is useful for making  
 diagnostic reagents or targeted enzymes that have therapeutic (*e.g.*, TEPT) or other

WO 03/107009 PCT/US03/18187  
applications. *See, e.g.*, U.S. Pat. App. Ser. No.s 10/022,073 and 10/022,097, both filed  
December 13, 2001, incorporated herein by reference in their entireties.

In another embodiment, the reporter fusion is made by grafting one or more  
binding sequences into a reporter sequence, or by grafting one or more reporter sequences into  
a binding sequence, *e.g.*, as described in copending U.S. Pat. App. Ser. No.s 10/022,073 and  
10/022,097, both filed December 13, 2001, or in copending U.S. Pat. App. Ser. No.  
60,279,609 and United States Serial Number 10,170,387 (attorney docket no. 9342-041 and  
40-999), incorporated herein by reference in their entireties.

### Binding Molecules

A prototype binding molecule comprises a molecule that has a measurable  
binding affinity for a target of interest. The prototype binding molecule can be any type of  
molecule, for example, a small organic molecule, a biological molecule (*e.g.*, a peptide, a  
polypeptide, a protein, a nucleic acid, an oligonucleotide, a polynucleotide, a sugar, a  
metabolite, a lipid, a vitamin, a co-factor, a nucleotide or an amino acid), a polymer, a drug, or  
an inorganic molecule. In one embodiment, the prototype binding molecule is a prototype  
binding sequence. A prototype binding sequence comprises a peptide, either alone or  
covalently attached to one or more other molecules, that binds to a target. The peptide can  
have any amino acid sequence and can have one or more covalent modifications. In one  
embodiment, the prototype binding sequence is an antibody, antibody fragment, or derivative.  
In another embodiment, the prototype binding sequence is not an antibody, antibody  
fragment, or derivative.

A variant binding molecule is similar to a prototype binding molecule that  
binds a target but differs from it in one or more aspects. The difference can be any difference  
that affects a binding property of the binding molecule. The difference can be, for example,  
one or more insertions, deletions and/or substitutions, or combinations thereof, of atoms,  
amino acids or functional groups. In one embodiment, the binding molecule is a binding  
sequence, and the difference is a difference in the amino acid sequence of the prototype  
binding sequence. The variant binding sequence can be, for example, at least 50, 55, 60, 65,  
70, 75, 80, 85, 90, 95, 98 or 99% identical to the prototype binding sequence. The amino  
acid sequence of the variant binding sequence can differ from the amino acid sequence of the  
prototype binding sequence by the presence or absence of one or more non-classical amino  
acids or chemical amino acid analogs. Non-classical amino acids include but are not limited  
to the D-isomers of the common amino acids,  $\alpha$ -amino isobutyric acid, 4-aminobutyric acid



(4-Abu), 2-aminobutyric acid (2- Abu), 6-amino hexanoic acid (Ahx), 2-amino isobutyric acid (2-Aib), 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine,  $\beta$ -alanine, fluoro-amino acids, designer amino acids such as  $\beta$ -methyl amino acids,  $\text{Ca}$ -methyl amino acids,  $\text{Na}$ -methyl amino acids, and amino acid analogs in general.

In another embodiment, the variant binding sequences has or lacks a covalent modification relative to the prototype binding sequence, for example, glycosylation, methylation, acetylation, phosphorylation, amidation, derivatization by protecting/blocking groups, proteolytic cleavage, *etc.*, as well as any of other numerous chemical modifications, including but not limited to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease,  $\text{NaBH}_4$ , acetylation, formylation, oxidation, reduction, metabolic synthesis in the presence of tunicamycin, *etc.*

These variant sequences can be rationally designed or can be generated by random or semi-random insertions, deletions or substitutions.

The binding molecules of the invention can bind a target with any affinity, *e.g.*, with a  $K_d$  of about 100  $\mu\text{M}$  or less, 10  $\mu\text{M}$  or less, 1  $\mu\text{M}$  or less, 100 nM or less, about 90 nM or less, about 80 nM or less, about 70 nM or less, about 60 nM or less, about 50 nM or less, about 40 nM or less, about 30 nM or less, about 20 nM or less, about 10 nM or less, about 5 nM or less, about 1 nM or less or about 0.1 nM or less. In one embodiment, the affinity, or another binding characteristic, of a binding molecule for a target is dependent on the conditions under which the binding is conducted. Examples of conditions affecting binding include pH, temperature, light, oxygen tension, salt concentration, the presence or absence of binding co-factors, and other conditions found in copending U.S. Pat. App. Ser. No. 60/279,609 (attorney docket no. 9342-042-999), filed concurrently with the present application, incorporated herein by reference in its entirety.

In one embodiment, the binding molecule is or is part of a targeted enzyme, *e.g.*, as described in copending U.S. Pat. App. Ser. No.s 10/022,073 and 10/022,097, both filed December 13, 2001, incorporated herein by reference in their entireties.

In another embodiment, the binding molecule is or is part of a milieu-dependent binding molecule, *e.g.*, a milieu-dependent targeted agent as described in copending U.S. Pat. App. Ser. No. 60/388,387 (attorney docket no. 9342-042-999), filed concurrently with the present application, incorporated herein by reference in its entirety.

In another embodiment, the binding molecule is or is part of a multifunctional polypeptide, *e.g.*, as described in copending U.S. Pat. App. Ser. No. 10/170,387 (attorney

### Reporter molecules

5 A reporter molecule can be any molecule that can be detected without the  
necessity of being bound by a detectable molecule, *e.g.*, a labeled antibody or other type of  
peptide or molecule that is labeled and binds to the reporter molecule. The reporter molecule  
additionally can have one or more desirable traits, for example, sensitive detection, selection  
10 of clones that produce a reporter fusion comprising the reporter molecule and a binding  
molecule of interest, stabilization of the reporter fusion or the binding molecule, protease-  
resistance of the reporter fusion or the binding sequence, easy purification, good expression or  
secretion of product into culture medium. Examples of reporter molecules include  
radiolabeled substances, fluorescent molecules, light-emitting molecules and molecules  
15 catalyzing or otherwise participating in a detectable chemical reaction, *e.g.*, a colorimetric  
reaction.

In one embodiment, the reporter molecule is a reporter sequence. In another  
embodiment, the reporter sequence is an enzyme. The enzyme can be any enzyme, or  
fragment or derivative of an enzyme, that can catalyze the transformation of a substrate into a  
detectable reaction product. Examples of enzymes that can be used as reporter sequences  
20 include  $\beta$ -lactamases,  $\beta$ -galactosidases, phosphatases, peroxidases, reductases, esterases,  
hydrolases, isomerases and proteases.

In one embodiment, the reporter sequence is the enzyme  $\beta$ -lactamase (BLA).  
The enzyme is highly active towards the specific substrate nitrocefin which allows the  
detection of bound reporter fusions at very low concentrations. One can synthesize substrates  
25 with higher sensitivity by using fluorogenic leaving groups. These substrates can be designed  
in analogy to BLA-activated prodrugs. *See, e.g.*, Hudyma *et al.*, 1993, Bioorg Med Chem  
Lett 3:323-28. BLA can be expressed in high concentration in *E.coli*. In one embodiment,  
the present invention provides expression vectors that release substantial amounts of BLA into  
the culture medium which greatly simplifies screening. BLA confers antibiotic resistance to  
30 its host. This can be exploited to quickly evaluate the success of a cloning experiment. One  
can attach a binding sequence to the N-terminus or C-terminus of BLA. If the mutagenesis of  
the binding sequence leads to sequences that are not correctly translated or that interfere with  
the cell physiology then such undesirable mutants can in be rapidly identified or eliminated by  
selection with an appropriate antibiotic like cefotaxime or carbenicillin.

BLA and its fusion products can be easily purified by affinity chromatography using immobilized phenylboronic acid or similar inhibitors. *See, e.g., Cartwright et al., 1984, Biochem J 221:505-12.*

Using prodrugs that have been developed for cancer treatment it is possible to select for cells that do not express BLA activity. This can be used to identify mutants where the BLA gene has been inactivated.

In another embodiment, the BLA has a specific activity greater than about 0.01 U/pmol against nitrocefin using the assay described in United States Patent Application Serial Number 10/022,097, filed December 13, 2001, incorporated herein by reference in its entirety.

In another embodiment, the specific activity is greater than about 0.1 U/pmol. In another embodiment, the specific activity is greater than about 1 U/pmol.

BLA enzymes are widely distributed in both gram-negative and gram-positive bacteria. BLA sequences are well known. A representative example of a BLA sequence is depicted in Figure 7. BLA enzymes vary in specificity, but have in common that they hydrolyze  $\beta$ -lactams, producing substituted  $\beta$ -amino acids. Thus, they confer resistance to antibiotics containing  $\beta$ -lactams. Because BLA enzymes are not endogenous to mammals, they are subject to minimal interference from inhibitors, enzyme substrates, or endogenous enzyme systems and therefore are particularly well-suited for therapeutic administration. BLA enzymes are further well-suited to the therapeutic methods of the present invention because of their small size (BLA from *E. cloacae* is a monomer of 43 kD; BLA from *E. coli* is a monomer of 30 kD) and because they have a high specific activity against their substrates and have optimal activity at 37° C. *See Melton et al., Enzyme-Prodrug Strategies for Cancer Therapy, Kluwer Academic/Plenum Publishers, New York (1999).*

The  $\beta$ -lactamases have been divided into four classes based on their sequences. *See Thomson et al., 2000, Microbes and Infection 2:1225-35.* The serine  $\beta$ -lactamases are subdivided into three classes: A (penicillinases), C (cephalosporinases) and D (oxacillinases). Class B  $\beta$ -lactamases are the zinc-containing or metallo  $\beta$ -lactamases. Any class of BLA can be utilized to generate reporter sequence of the invention.

In one embodiment, the present invention provides a BLA reporter sequence that comprises the sequence YXN at its substrate recognition site (throughout, "X" refers to any amino acid residue). In another embodiment, the BLA reporter sequence comprises the sequence RLYANASI at its active site. In another embodiment, the BLA reporter sequence comprises a sequence at its active site that differs from the sequence RLYANASI by one, two or three amino acid residues. The differences can be, for example, the substitution of

conservative amino acid residues, insertions, deletions and non-conservative amino acid substitutions.

In another embodiment, the present invention provides a BLA reporter sequence that comprises the sequence KTXS at its substrate recognition site. In another embodiment, the BLA reporter sequence comprises the sequence VHKTGSTG at its active site. In another embodiment, the BLA reporter sequence comprises a sequence at its active site that differs from the sequence VHKTGSTG by one, two or three amino acid residues. The differences can be, for example, the substitution of conservative amino acid residues, insertions, deletions and non-conservative amino acid substitutions.

In another embodiment, the present invention provides a BLA reporter sequence that comprises the sequences YXN and KTXS at its substrate recognition site. In another embodiment, the BLA reporter sequence comprises the sequences VHKTGSTG and RLYANASI at its active site. In another embodiment, the BLA reporter sequence comprises sequences at its active site that differ from the sequences RLYANASI and VHKTGSTG by one, two or three amino acid residues. The differences can be, for example, the substitution of conservative amino acid residues, insertions, deletions and non-conservative amino acid substitutions.

In one embodiment, the BLA reporter sequence comprises the amino acid sequence of Figure 7. In another embodiment, the BLA reporter sequence is at least 50%, 60%, 70%, 80%, 90%, 95%, 98 or 99% or more identical to the sequence depicted in Figure 7.

In another embodiment, a nucleic acid encoding the BLA reporter sequence hybridizes to a nucleic acid complementary to a nucleic acid encoding the amino acid sequence of Figure 7 under highly stringent conditions. The highly stringent conditions can be, for example, hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65° C, and washing in 0.1xSSC/0.1 % SDS at 68° C (Ausubel *et al.*, eds., 1989, *Current Protocols in Molecular Biology*, Vol. I, Green Publishing Associates, Inc., and John Wiley & Sons, Inc., New York, at p. 2.10.3). Other highly stringent conditions can be found in, for example, *Current Protocols in Molecular Biology*, at pages 2.10.1-16 and *Molecular Cloning: A Laboratory Manual*, 2d ed., Sambrook *et al.* (eds.), Cold Spring Harbor Laboratory Press, 1989, pages 9.47-57. In another embodiment, a nucleic acid encoding the BLA reporter sequence hybridizes to a nucleic acid complementary to a nucleic acid encoding the amino acid sequence of Figure 7 under moderately stringent conditions. The moderately stringent conditions can be, for example, washing in

0.2xSSC/0.1% SDS at 42° C (Ausubel et al., 1989, *supra*). Other moderately stringent conditions can be found in, for example, *Current Protocols in Molecular Biology*, Vol. I, Ausubel et al. (eds.), Green Publishing Associates, Inc., and John Wiley & Sons, Inc., 1989, pages 2.10.1-16 and *Molecular Cloning: A Laboratory Manual*, 2d ed., Sambrook et al.

5 (eds.), Cold Spring Harbor Laboratory Press, 1989, pages 9.47-57.

Fluorescent reporters like green fluorescent protein (GFP) or red fluorescent protein (RFP) also can be used.

In one embodiment, the reporter fusion comprises a plurality of reporter molecules, e.g., a plurality of reporter sequences. In a particular embodiment, a reporter  
10 fusion comprises a reporter sequence at its N-terminus and at its C-terminus. A reporter fusion comprising a plurality of reporter sequences is particularly useful if the goal is to screen for protease-resistant variants of a binding sequence. In one embodiment, the reporter fusion comprises a BLA reporter sequence and a fluorescent reporter sequence, e.g., GFP or RFP. The BLA reporter can be used for antibiotic selection and purification and the GFP  
15 reporter for detection (e.g., using FACS).

### Targets

The targets of the present invention can be any substance or composition to which a molecule can be made to bind.

20 In one aspect, the target is a surface. In one embodiment, the surface is a biological surface. In another embodiment, the biological surface is a surface of an organ. In another embodiment, the biological surface is a surface of a tissue. In another embodiment, the biological surface is a surface of a cell. In another embodiment, the biological surface is a surface of a diseased organ, tissue or cell. In another embodiment, the biological surface is a  
25 surface of a normal or healthy organ, tissue or cell. In another embodiment, the surface is a macromolecule in the interstitial space of a tissue. In another embodiment, the biological surface is the surface of a virus or pathogen. In another embodiment, the surface is a non-biological surface. In another embodiment, the non-biological surface is a surface of a medical device. In another embodiment, the medical device is a therapeutic device. In  
30 another embodiment, the therapeutic device is an implanted therapeutic device. In another embodiment, the medical device is a diagnostic device. In another embodiment, the diagnostic device is a well or tray.

Sources of cells or tissues include human, animal, bacterial, fungal, viral and plant. Tissues are complex targets and refer to a single cell type, a collection of cell types or

an aggregate of cells generally of a particular kind. Tissue may be intact or modified. General classes of tissue in humans include but are not limited to epithelial tissue, connective tissue, nerve tissue, and muscle tissue.

In another aspect, the target is a cancer-related target. The cancer-related target can be any target that a composition of the invention binds to as part of the diagnosis, detection or treatment of a cancer or cancer-associated condition in a subject, for example, a cancerous cell, tissue or organ, a molecule associated with a cancerous cell, tissue or organ, or a molecule, cell, tissue or organ that is associated with a cancerous cell, tissue or organ (e.g., a tumor-bound diagnostic or therapeutic molecule administered to a subject or to a biopsy taken from a subject, or a healthy tissue, such as vasculature, that is associated with cancerous tissue). Examples of cancer-related targets are provided in U.S. Pat. No. 6,261,535, which is incorporated herein by reference in its entirety.

The cancer-related target can be related to any cancer or cancer-associated condition. Examples of types of cancers include carcinomas, sarcomas, myelomas, leukemias, lymphomas and mixed type cancers.

In one embodiment, the cancer is a bone cancer, for example, Ewing's sarcoma, osteosarcoma and rhabdomyosarcoma and other soft-tissue sarcomas. In another embodiment, the cancer is a brain tumor, for example, oligodendroglioma, ependymoma, meningioma, lymphoma, schwannoma or medulloblastoma. In another embodiment, the cancer is breast cancer, for example, ductal carcinoma *in situ* of the breast. In another embodiment, the cancer is an endocrine system cancer, for example, adrenal, pancreatic, parathyroid, pituitary and thyroid cancers. In another embodiment, the cancer is a gastrointestinal cancer, for example, anal, colorectal, esophageal, gallbladder, gastric, liver, pancreatic, and small intestine cancers. In another embodiment, the cancer is a gynecological cancer, for example, cervical, endometrial, uterine, fallopian tube, gestational trophoblastic disease, choriocarcinoma, ovarian, vaginal, and vulvar cancers. In another embodiment, the cancer is a head and neck cancer, for example, laryngeal, oropharyngeal, parathyroid or thyroid cancer. In another embodiment, the cancer is a leukemic cancer, for example, acute lymphocytic leukemia, acute myelogenous leukemia, chronic lymphocytic leukemia, chronic myelogenous leukemia, hairy cell leukemia, or a myeloproliferative disorder. In another embodiment, the cancer is a lung cancer, for example, a mesothelioma, non-small cell small cell lung cancer. In another embodiment, the cancer is a lymphoma, for example, AIDS-related lymphoma, cutaneous T cell lymphoma/mucosis fungoides, Hodgkin's disease, or non-Hodgkin's disease. In another embodiment, the cancer is metastatic cancer. In another

embodiment, the cancer is a myeloma, for example, a multiple myeloma. In another embodiment, the cancer is a pediatric cancer, for example, a brain tumor, Ewing's sarcoma, leukemia (*e.g.*, acute lymphocytic leukemia or acute myelogenous leukemia), liver cancer, a lymphoma (*e.g.*, Hodgkin's lymphoma or non-Hodgkin's lymphoma), neuroblastoma, retinoblastoma, a sarcoma (*e.g.*, osteosarcoma or rhabdomyosarcoma), or Wilms' Tumor. In another embodiment, the cancer is penile cancer. In another embodiment, the cancer is prostate cancer. In another embodiment, the cancer is a sarcoma, for example, Ewing's sarcoma, osteosarcoma, rhabdomyosarcoma and other soft-tissue sarcomas. In another embodiment, the cancer is a skin cancer, for example, cutaneous T cell lymphoma, mycosis fungoides, Kaposi's sarcoma or melanoma. In another embodiment, the cancer is testicular cancer. In another embodiment, the cancer is thyroid cancer, for example, papillary, follicular, medullary, or anaplastic or undifferentiated thyroid carcinoma. In another embodiment, the cancer is urinary tract cancers, for example, bladder, kidney or urethral cancers. In another embodiment, the cancer or cancer-related condition is ataxia-telangiectasia, carcinoma of unknown primary origin, Li-Fraumeni syndrome, or thymoma.

In another aspect, the cancer-related target is a molecule associated with a cancerous cell or tissue. In one embodiment, the molecule is a tumor or tumor stroma antigen, for example, GD2, Lewis-Y, 72 kd glycoprotein (gp72, decay-accelerating factor, CD55, DAF, C3/C5 convertases), CO17-1A (EpCAM, 17-1A, EGP-40), TAG-72, CSAg-P (CSAp), 45kd glycoprotein, HT-29 ag, NG2, A33 (43kd gp), 38kd gp, MUC-1, CEA, EGFR (HER1), HER2, HER3, HER4, HN-1 ligand, CA125, syndecan-1, Lewis X, PgP, FAP stromal Ag (fibroblast activation protein), EDG receptors (endoglin receptors), ED-B, laminin-5 (gamma2), cox-2 (+LN-5), PgP (P-glycoprotein), alphaVbeta3 integrin, alphaVbeta5, integrin, uPAR (urokinase plasminogen activator receptor), endoglin (CD105), folate receptor osteopontin (EDG 1,3), p97 (melanotransferrin), farnesyl transferase or a molecule in an apoptotic pathway (*e.g.*, a death receptor, fas, caspase or bcl-2) or a lectin.

In another aspect, the target is a hematopoietic cell. Hematopoietic cells encompass hematopoietic stem cells (HSCs), erythrocytes, neutrophils, monocytes, platelets, mast cells, eosinophils, basophils, B and T cells, macrophages, and natural killer cells. In one embodiment, the HSC has a surface antigen expression profile of CD34<sup>+</sup> Thy-1<sup>+</sup>, and preferably CD34<sup>+</sup> Thy-1<sup>+</sup> Lin<sup>-</sup>. Lin<sup>-</sup> refers to a cell population selected on the basis of the lack of expression of at least one lineage specific marker. Methods for isolating and selecting HSCs are well known in the art and reference is made to U.S. Patent Nos. 5,061,620, 5,677,136, and 5,750,397, each of which is incorporated herein in its entirety.

In another aspect, the target is a molecule. In one embodiment, the molecule is an organic molecule. In another embodiment, the molecule is a biological molecule. In another embodiment, the biological molecule is a cell-associated molecule. In another embodiment, the cell-associated molecule is associated with the outer surface of a cell. In another embodiment, the cell-associated molecule is part of the extracellular matrix. In another embodiment, the cell-associated molecule is associated with the outer surface of a cell is a protein. In another embodiment, the protein is a receptor. In another embodiment, the cell-associated molecule is specific to a type of cell in a subject. In another embodiment, the type of cell is a diseased cell. In another embodiment, the diseased cell is a cancer cell. In another embodiment, the diseased cell is an infected cell. Other molecules that can serve as targets according to the invention include, but are not limited to, proteins, peptides, nucleic acids, carbohydrates, lipids, polysaccharides, glycoproteins, hormones, receptors, antigens, antibodies, toxic substances, metabolites, inhibitors, drugs, dyes, nutrients and growth factors.

In another aspect, the target is a surface feature, the surface feature comprising two or more molecules. The two or more molecule may include, but are not limited to, proteins, peptides, nucleic acids, carbohydrates, lipids, polysaccharides, glycoproteins, hormones, receptors, antigens, antibodies, toxic substances, metabolites, inhibitors, drugs, dyes, nutrients or growth factors.

Non-limiting examples of protein and chemical targets encompassed by the invention include chemokines and cytokines and their receptors. Cytokines as used herein refer to any one of the numerous factors that exert a variety of effects on cells, for example inducing growth or proliferation. Non-limiting examples include interleukins (IL), IL-2, IL-3, IL-4 IL-6, IL-10, IL-12, IL-13, IL-14 and IL-16; soluble IL-2 receptor; soluble IL-6 receptor; erythropoietin (EPO); thrombopoietin (TPO); granulocyte macrophage colony stimulating factor (GM-CSF); stem cell factor (SCF); leukemia inhibitory factor (LIF); interferons; oncostatin M (OM); the immunoglobulin superfamily; tumor necrosis factor (TNF) family, particularly TNF- $\alpha$ ; TGF $\beta$ ; and IL-1 $\alpha$ ; and vascular endothelial growth factor (VEGF) family, particularly VEGF (also referred to in the art as VEGF-A), VEGF-B, VEGF-C, VEGF-D and placental growth factor (PLGF). Cytokines are commercially available from several vendors including Amgen (Thousand Oaks, CA), Immunex (Seattle, WA) and Genentech (South San Francisco, CA). Particularly preferred are VEGF and TNF- $\alpha$ . Antibodies against TNF- $\alpha$  show that blocking interaction of the TNF- $\alpha$  with its receptor is useful in modulating over-expression of TNF- $\alpha$  in several disease states such as septic shock, rheumatoid arthritis, or other inflammatory processes. VEGF is an angiogenic inducer, a mediator of vascular



permeability, and an endothelial cell specific mitogen. VEGF has also been implicated in tumors. Targeting members of the VEGF family and their receptors may have significant therapeutic applications, for example blocking VEGF may have therapeutic value in ovarian hyper stimulation syndrome (OHSS). Reference is made to N. Ferrara *et al.*, (1999) *Nat.*

5 *Med.* 5:1359 and Gerber *et al.*, (1999) *Nat. Med.* 5:623. Other preferred targets include cell-surface receptors, such as T-cell receptors.

Chemokines are a family of small proteins that play an important role in cell trafficking and inflammation. Members of the chemokine family include, but are not limited to, IL-8, stomal-derived factor-1(SDF-1), platelet factor 4, neutrophil activating protein-2

10 (NAP-2) and monocyte chemo attractant protein-1 (MCP-1).

Other protein and chemical targets include, but are not limited to:

immunoregulation modulating proteins, such as soluble human leukocyte antigen (HLA, class I and/or class II, and non-classical class I HLA (E, F and G)); surface proteins, such as soluble T or B cell surface proteins; human serum albumin; arachadonic acid metabolites, such as

15 prostaglandins, leukotrienes, thromboxane and prostacyclin; IgE, auto or alloantibodies for autoimmunity or allo- or xenoimmunity, Ig Fc receptors or Fc receptor binding factors; G-protein coupled receptors; cell-surface carbohydrates; angiogenesis factors; adhesion molecules; ions, such as calcium, potassium, magnesium, aluminum, and iron; fibril proteins, such as prions and tubulin; enzymes, such as proteases, aminopeptidases, kinases,

20 phosphatases, DNAses, RNAases, lipases, esterases, dehydrogenases, oxidases, hydrolases, sulphatases, cyclases, transferases, transaminases, carboxylases, decarboxylases, superoxide dismutase, and their natural substrates or analogs; hormones and their corresponding receptors, such as follicle stimulating hormone (FSH), leutinizing hormone (LH), thyroxine (T4 and T3), apolipoproteins, low density lipoprotein (LDL), very low density lipoprotein

25 (VLDL), cortisol, aldosterone, estriol, estradiol, progesterone, testosterone,

dehydroepiandrosterone (DHBA) and its sulfate (DHEA-S); peptide hormones, such as renin, insulin, calcitonin, parathyroid hormone (PTH), human growth hormone (hGH), vasopressin and antidiuretic hormone (AD), prolactin, adrenocorticotrophic hormone (ACTH), LHRH, thyrotropin-releasing hormone (TRH), vasoactive intestinal peptide (VIP), bradykinin and

30 corresponding prohormones; catecholamines such as adrenaline and metabolites; cofactors including atrionatriutic factor (AdF), vitamins A, B, C, D, E and K, and serotonin; coagulation factors, such as prothrombin, thrombin, fibrin, fibrinogen, Factor VIII, Factor IX, Factor XI, and von Willebrand factor; plasminogen factors, such as plasmin, complement activation factors, LDL and ligands thereof, and uric acid; compounds regulating coagulation, such as

hirudin, hirulog, hementin, heparin, and tissue plasminogen activator (TPA); nucleic acids for gene therapy; compounds which are enzyme antagonists; and compounds binding ligands, such as inflammation factors; and receptors and other proteins that bind to one or more of the preceding molecules.

5 Non-human derived targets include without limitation drugs, especially drugs subject to abuse, such as cannabis, heroin and other opiates, phencyclidine (PCP), barbiturates, cocaine and its derivatives, and benzodiazepine; toxins, such as heavy metals like mercury and lead, arsenic, and radioactive compounds; chemotherapeutic agents, such as paracetamol, digoxin, and free radicals; bacterial toxins, such as lipopolysaccharides (LPS)  
10 and other gram negative toxins, *Staphylococcus* toxins, Toxin A, Tetanus toxins, Diphtheria toxin and Pertussis toxins; plant and marine toxins; snake and other venoms, virulence factors, such as aerobactins, or pathogenic microbes; infectious viruses, such as hepatitis, cytomegalovirus (CMV), herpes simplex virus (HSV types 1, 2 and 6), Epstein-Barr virus (EBV), varicella zoster virus (VZV), human immunodeficiency virus (HIV-1, -2) and other  
15 retroviruses, adenovirus, rotavirus, influenzae, rhinovirus, parvovirus, rubella, measles, polio, pararyxovirus, papovavirus, poxvirus and picornavirus, prions, plasmodia tissue factor, protozoans, such as *Entamoeba histolitica*, Filaria, Giardia, Kalaazar, and toxoplasma; bacteria, gram-negative bacteria responsible for sepsis and nosocomial infections such as *E. coli*, *Acynetobacter*, *Pseudomonas*, *Proteus* and *Klebsiella*, also gram-positive bacteria such  
20 as *Staphylococcus*, *Streptococcus*, *Meningococcus* and *Llycobacteria*, *Chlamydiae* *Legionnella* and Anaerobes; fungi such as *Candida*, *Pneumocystis*, *Aspergillus*, and *Mycoplasma*.

In one aspect the target includes an enzyme such as proteases, aminopeptidases, kinases, phosphatases, DNAses, RNAases, lipases, esterases,  
25 dehydrogenases, oxidases, hydrolases, sulphatases, cellulases, cyclases, transferases, transaminases, carboxylases, decarboxylases, superoxide dismutase, and their natural substrates or analogs. Particularly preferred enzymes include hydrolases, particularly alpha/beta hydrolases; serine proteases, such as subtilisins, and chymotrypsin serine proteases; cellulases; and lipases.

30 In another embodiment, the target is a non-biological material. In another embodiment, the non-biological material is a fabric. In another embodiment, the fabric is a natural fabric. In another embodiment, the fabric is cotton. In another embodiment, the fabric is silk. In another embodiment, the fabric is wool. In another embodiment, the fabric is a non-natural fabric. In another embodiment, the fabric is nylon. In another embodiment, the

fabric is rayon. In another embodiment, the fabric is polyester. In another embodiment, the non-biological material is a plastic. In another embodiment, the non-biological material is a ceramic. In another embodiment, the non-biological material is a metal. In another embodiment, the non-biological material is rubber. In another embodiment, the non-biological material is wood.

In another embodiment the target is a microcircuit. This circuit can be in its finished form or in any stage of circuit manufacturing. *See, e.g., van Zant, 2000, Microchip Fabrication*, McGraw-Hill, New York, incorporated herein by reference in its entirety.

In another embodiment, the target is not an antibody (*e.g., a polyclonal antibody, a monoclonal antibody, an scFv, or another antigen-binding fragment of an antibody*).

#### Methods of Selecting Variant Binding Molecules

In one aspect, the present invention provides methods of screening reporter fusions comprising variant binding molecules and reporter molecules to identify binding molecules with desired binding characteristics. Any method of screening reporter fusions comprising variant binding molecules and reporter molecules that can identify binding molecules with improved binding characteristics can be used.

Reporter fusions can be used in various ways that allow one to assay for different properties of the binding molecule. For instance, by allowing for sufficient time between measurements to reach binding equilibrium one can identify variants with improved binding affinity. Alternatively, one can screen a population of variants under two or more different conditions to identify variants that differentiate between various targets or variants that show differential binding in dependence of the reaction conditions. *See, e.g., U.S. Pat. App. Ser. No. 60/388,387 (attorney docket no. 9342-042-999)*, filed concurrently with the present application, incorporated herein by reference in its entirety.

In one embodiment, the binding molecule is a binding sequence. One embodiment of a screen for selecting an improved binding sequence is illustrated in Figure 3. The process starts with a population of clones, where various clones produce a different reporter fusion comprising a different variant of a prototype binding sequence. The clones are cultured under conditions that facilitate the expression of the reporter fusions. In one embodiment, reporter fusions are released by the clones into the culture medium. Subsequently, a part of the culture is transferred to a microtiter plate to which the target of interest has been bound. After incubation to allow target-reporter fusion interaction, unbound

reporter fusion is removed, for example, by washing or filtration. Then a chromogenic substrate is added to determine the quantity of bound reporter fusion for each variant. During this measurement, a fraction of the bound reporter fusion can dissociate from the target. Subsequently, one can remove the dissociated reporter fusion and add fresh substrate to measure the remaining concentration of bound reporter fusion. This process of washing and measuring can be repeated several times. As a result, one can determine the dissociation rate of each variant in the population and detect variants with improved binding properties.

Figure 5 shows, as an example, a screen at two different pH values. By comparing the values obtained under both pH conditions one can identify variants that show pH-dependent binding to their target. Further examples of methods of screening for binding sequences that bind to a target better under a first set of conditions than they bind under a second set of conditions are provided in copending U.S. Pat. App. Ser. No. 60/388,387 (attorney docket no. 9342-042-999), filed concurrently with the present application, incorporated herein by reference in its entirety. In a similar way, one can compare binding in the presence of different effector molecules to obtain variants which show effector-dependent binding to a target.

In another embodiment, an improved binding molecule is selected by contacting a non-biological target with the reporter fusion, for example, a computer chip at any stage during its manufacture, or after it is manufactured, or a surface (*e.g.*, glass, plastic, fabric, film or membrane) exhibiting, *e.g.*, coated with, a target molecule.

Several methods have been described for manufacturing arrays of compounds. These methods can be adapted to screen populations of reporter fusions for binding to a target. One embodiment of this process is illustrated in Figure 6. Aliquots of variants are transferred onto a surface (*e.g.*, membrane, plastic or glass) which carries bound target. The target can be, for example, a molecule or a cell of interest. Unbound reporter fusions are removed by washing and a substrate is added to detect the remaining bound reporter fusion. It is important to use a substrate that can be used to detect surface-bound reporter fusion. Such substrates are commonly used for immunohistochemistry, *e.g.*, 5-bromo-4-chloro-2-indoyl  $\beta$ -D-galactopyranoside, diaminobenzidine, ELF® 97 esterase substrate (Molecular Probes, Eugene, OR), ELF® 97 phosphatase substrate (Molecular Probes), ELF® 97  $\beta$ -D-glucuronide, ELF® 97 N-acetylglucosaminide.

In addition, one can use fluorescent reporters to screen for variant binding sequences that bind to a target of interest, *e.g.*, a cell. In one embodiment, the assay comprises one or more of the following steps:

Generating population of reporter fusions in a suitable host;  
Growing host clones to produce reporter fusions;  
Mixing the reporter fusions with a cell suspension;  
Adding a fluorescent reference protein that shows fluorescence that can be  
distinguished from the fluorescent reporter; and  
Analyzing each cell suspension in a fluorescence activated cell sorter (FACS)  
to identify clones of reporter fusions that differ in their binding behavior from  
the control protein.

In one embodiment, a variant binding sequence is selected by immobilizing  
reporter fusion-producing cells in agarose beads or similar material to which the target has  
been attached. The reporter fusion can bind to the target in the bead and it can be detected by,  
for example, using a fluoregenic substrate, which allows stained beads to be sorted using a  
fluorescence-activated cell sorter (FACS). *See, e.g., Gray et al., 1995, J. Immun. Meth.*  
182:155-63.

In another embodiment, the screening method comprises multiple rounds of  
generating a variant binding molecule of a prototype binding molecule, contacting the target  
with a reporter fusion comprising the variant binding molecule and a reporter molecule, and  
selecting the variant sequence if it binds to the target with a desired binding characteristic,  
wherein the variant binding molecule selected in a previous selection step is the prototype  
binding molecule of the subsequent generating step. Using this approach, multiple rounds of  
screening can be used to select binding molecules with increasingly refined binding  
characteristics.

In another embodiment, the screening method comprises contacting the target  
with a library of reporter fusions, wherein various reporter fusions in the library comprise a  
different variant binding molecules. In a more particularly defined embodiment, the method  
comprises using multiple rounds of screening, as described herein, wherein in each round the  
target is contacted with a library of reporter fusions comprising variant binding molecules that  
are derived from the binding molecule selected in a previous round.

#### Nucleic Acids and Methods of Making Reporter Sequences, Binding Sequences and Reporter Fusions

In another aspect, the present invention provides a nucleic acid encoding a  
polypeptide comprising all or part of a reporter sequence, a binding sequence or a reporter  
fusion. The nucleic acid can be, for example, a DNA or an RNA. The present invention also

provides a plasmid comprising a nucleic acid encoding a polypeptide comprising all or part of a reporter sequence, a binding sequence or a reporter fusion. The plasmid can be, for example, an expression plasmid that allows expression of the polypeptide in a host cell or organism, or *in vitro*. The expression vector can allow expression of the polypeptide in, for example, a bacterial cell. The bacterial cell can be, for example, an *E. coli* cell.

Because of the redundancy in the genetic code, typically a large number of DNA sequences encode any given amino acid sequence and are, in this sense, equivalent. As described below, it may be desirable to select one or another equivalent DNA sequences for use in a expression vector, based on the preferred codon usage of the host cell into which the expression vector will be inserted. The present invention is intended to encompass all DNA sequences that encode the reporter sequence, binding sequence or reporter fusion.

Production of the reporter sequence, binding sequence or reporter fusion of the invention can be carried out using a recombinant expression clone. The construction of the recombinant expression clone, the transformation of a host cell with the expression clone, and the culture of the transformed host cell under conditions which promote expression, can be carried out in a variety of ways using techniques of molecular biology well understood in the art. Methods for each of these steps are described in general below. Preferred methods are described in detail in the examples.

An operable expression clone is constructed by placing the coding sequence in operable linkage with a suitable control sequences in an expression vector. The vector can be designed to replicate autonomously in the host cell or to integrate into the chromosomal DNA of the host cell. The resulting clone is used to transform a suitable host, and the transformed host is cultured under conditions suitable for expression of the coding sequence. The expressed reporter sequence, binding sequence or reporter fusion is isolated from the medium or from the cells, although recovery and purification of the reporter sequence, binding sequence or reporter fusion may not be necessary in some instances.

Construction of suitable clones containing the coding sequence and a suitable control sequence employs standard ligation and restriction techniques that are well understood in the art. In general, isolated plasmids, DNA sequences, or synthesized oligonucleotides are cleaved, modified, and religated in the form desired. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to facilitate construction of an expression clone.

Site-specific DNA cleavage is performed by treating with a suitable restriction enzyme (or enzymes) under conditions that are generally understood in the art and specified

by the manufacturers of commercially available restriction enzymes. *See, e.g.,* product catalogs from Amersham (Arlington Heights, IL), Roche Molecular Biochemicals (Indianapolis, IN), and New England Biolabs (Beverly, MA). In general, about 1 µg of plasmid or other DNA is cleaved by one unit of enzyme in about 20 µl of buffer solution; in the examples below, an excess of restriction enzyme is generally used to ensure complete digestion of the DNA. Incubation times of about one to two hours at a temperature which is optimal for the particular enzyme are typical. After each incubation, protein is removed by extraction with phenol and chloroform; this extraction can be followed by ether extraction and recovery of the DNA from aqueous fractions by precipitation with ethanol. If desired, size separation of the cleaved fragments may be performed by polyacrylamide gel or agarose gel electrophoresis using standard techniques. *See, e.g.,* Maxam *et al.*, 1980, *Methods in Enzymology* 65:499-560.

Restriction enzyme-cleaved DNA fragments with single-strand "overhanging" termini can be made blunt-ended (double-strand ends) by, for example, treating with the large fragment of *E. coli* DNA polymerase I (Klenow) in the presence of the four deoxynucleoside triphosphates (dNTPs) using incubation times of about 15 to 25 minutes at 20° C to 25° C in 50 mM Tris, pH 7.6, 50 mM NaCl, 10 mM MgCl<sub>2</sub>, 10 mM DTT, and 5 to 10 µM dNTPs. The Klenow fragment fills in at 5' protruding ends, but chews back protruding 3' single strands, even though the four dNTPs are present. If desired, selective repair can be performed by supplying one or more selected dNTPs, within the limitations dictated by the nature of the protruding ends. After treatment with Klenow, the mixture is extracted with phenol/chloroform and ethanol precipitated. Similar results can be achieved using S1 nuclease, because treatment under appropriate conditions with S1 nuclease results in hydrolysis of any single-stranded portion of a nucleic acid.

Ligations can be performed, for example, in 15-30 µl volumes under the following standard conditions and temperatures: 20 mM Tris-Cl, pH 7.5, 10 mM MgCl<sub>2</sub>, 10 mM DTT, 33 µg/ml BSA, 10-50 mM NaCl, and either 40 µM ATP and 0.01-0.02 (Weiss) units T4 DNA ligase at 0° C (for ligation of fragments with complementary single-stranded ends) or 1mM ATP and 0.3-0.6 units T4 DNA ligase at 14° C (for "blunt end" ligation). Intermolecular ligations of fragments with complementary ends are usually performed at 33-100 µg/ml total DNA concentrations (5-100 nM total ends concentration). Intermolecular blunt end ligations (usually employing a 20-30 fold molar excess of linkers, optionally) are performed at 1 µM total ends concentration.

In vector construction, the vector fragment is commonly treated with bacterial or calf intestinal alkaline phosphatase (BAP or CIAP) to remove the 5' phosphate and prevent religation and reconstruction of the vector. BAP and CIAP digestion conditions are well known in the art, and published protocols usually accompany the commercially available BAP and CIAP enzymes. To recover the nucleic acid fragments, the preparation is extracted with phenol-chloroform and ethanol precipitated to remove the phosphatase and purify the DNA. Alternatively, religation of unwanted vector fragments can be prevented by restriction enzyme digestion before or after ligation, if appropriate restriction sites are available.

Correct ligations for plasmid construction can be confirmed using any suitable method known in the art. For example, correct ligations for plasmid construction can be confirmed by first transforming a suitable host, such as *E. coli* strain DG101 (ATCC 47043) or *E. coli* strain DG116 (ATCC 53606), with the ligation mixture. Successful transformants are selected by ampicillin, tetracycline or other antibiotic resistance or sensitivity or by using other markers, depending on the mode of plasmid construction, as is understood in the art. Plasmids from the transformants are then prepared according to the method of Clewell *et al.*, 1969, Proc. Natl. Acad. Sci. USA 62:1159, optionally following chloramphenicol amplification. See Clewell, 1972, J. Bacteriol. 110:667. Alternatively, plasmid DNA can be prepared using the "Base-Acid" extraction method at page 11 of the Bethesda Research Laboratories publication Focus 5 (2), and very pure plasmid DNA can be obtained by replacing steps 12 through 17 of the protocol with CsCl/ethidium bromide ultracentrifugation of the DNA. As another alternative, a commercially available plasmid DNA isolation kit, e.g., HISPEED™, QIAFILTER™ and QIAGEN™ plasmid DNA isolation kits (Qiagen, Valencia CA) can be employed following the protocols supplied by the vendor. The isolated DNA can be analyzed by, for example, restriction enzyme digestion and/or sequenced by the dideoxy method of Sanger *et al.*, 1977, Proc. Natl. Acad. Sci. USA 74:5463, as further described by Messing *et al.*, 1981, Nuc. Acids Res. 9:309, or by the method of Maxam *et al.*, 1980, Methods in Enzymology 65:499.

The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene. Generally, procaryotic, yeast, insect, or mammalian cells are used as hosts. Procaryotic hosts are in general the most efficient and convenient for the production of recombinant proteins and are therefore preferred for the expression of the protein.

The procaryote most frequently used to express recombinant proteins is *E. coli*. However, microbial strains other than *E. coli* can also be used, such as bacilli, for example



*Bacillus subtilis*, various species of *Pseudomonas* and *Salmonella*, and other bacterial strains. In such procaryotic systems, plasmid vectors that contain replication sites and control sequences derived from the host or a species compatible with the host are typically used.

For expression of constructions under control of most bacterial promoters, *E. coli* K12 strain MM294, obtained from the *E. coli* Genetic Stock Center under GCSC #6135, can be used as the host. For expression vectors with the P<sub>L</sub>NRBS or P<sub>L</sub> T7RBS control sequence, *E. coli* K12 strain MC1000 lambda lysogen, N7N53cI857 SusP80, ATCC 39531, may be used. *E. coli* DG116, which was deposited with the ATCC (ATCC 53606) on April 7, 1987, and *E. coli* KB2, which was deposited with the ATCC (ATCC 53075) on March 29, 1985, are also useful host cells. For M13 phage recombinants, *E. coli* strains susceptible to phage infection, such as *E. coli* K12 strain DG98 (ATCC 39768), are employed. The DG98 strain was deposited with the ATCC on July 13, 1984.

For example, *E. coli* is typically transformed using derivatives of pBR322, described by Bolivar *et al.*, 1977, Gene 2:95. Plasmid pBR322 contains genes for ampicillin and tetracycline resistance. These drug resistance markers can be either retained or destroyed in constructing the desired vector and so help to detect the presence of a desired recombinant. Commonly used procaryotic control sequences, i.e., a promoter for transcription initiation, optionally with an operator, along with a ribosome binding site sequence, include the  $\beta$ -lactamase (penicillinase) and lactose (*lac*) promoter systems, *see* Chang *et al.*, 1977, Nature 198:1056, the tryptophan (*trp*) promoter system, *see* Goeddel *et al.*, 1980, Nuc. Acids Res. 8:4057, and the lambda-derived P<sub>L</sub> promoter, *see* Shimatake *et al.*, 1981, Nature 292:128, and gene N ribosome binding site (NRBS). A portable control system cassette is set forth in U.S. Patent No. 4,711,845, issued December 8, 1987. This cassette comprises a P<sub>L</sub> promoter operably linked to the NRBS in turn positioned upstream of a third DNA sequence having at least one restriction site that permits cleavage within six base pairs 3' of the NRBS sequence. Also useful is the phosphatase A (*phoA*) system described by Chang *et al.*, in European Patent Publication No. 196,864, published October 8, 1986. However, any available promoter system compatible with procaryotes can be used to construct a expression vector of the invention.

In addition to bacteria, eucaryotic microbes, such as yeast, can also be used as recombinant host cells. Laboratory strains of *Saccharomyces cerevisiae*, Baker's yeast, are most often used, although a number of other strains are commonly available. While vectors employing the two micron origin of replication are common, *see* Broach, 1983, Meth. Enz.

101:307, other plasmid vectors suitable for yeast expression are known. *See, e.g.,* Stinchcomb *et al.*, 1979, *Nature* 282:39; Tschempe *et al.*, 1980, *Gene* 10:157; and Clarke *et al.*, 1983, *Meth. Enz.* 101:300. Control sequences for yeast vectors include promoters for the synthesis of glycolytic enzymes. *See* Hess *et al.*, 1968, *J. Adv. Enzyme Reg.* 7:149; Holland *et al.*,  
5 1978, *Biotechnology* 17:4900; and Holland *et al.*, 1981, *J. Biol. Chem.* 256:1385. Additional promoters known in the art include the promoter for 3-phosphoglycerate kinase, *see* Hitzeman *et al.*, 1980, *J. Biol. Chem.* 255:2073, and those for other glycolytic enzymes, such as glyceraldehyde 3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate  
10 kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. Other promoters that have the additional advantage of transcription controlled by growth conditions are the promoter regions for alcohol dehydrogenase 2, isocytichrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and enzymes responsible for maltose and galactose utilization (Holland, *supra*).

15 Terminator sequences may also be used to enhance expression when placed at the 3' end of the coding sequence. Such terminators are found in the 3' untranslated region following the coding sequences in yeast-derived genes. Any vector containing a yeast-compatible promoter, origin of replication, and other control sequences is suitable for use in constructing yeast expression vectors.

20 The coding sequence can also be expressed in eucaryotic host cell cultures derived from multicellular organisms. *See, e.g.,* Tissue Culture, Academic Press, Cruz and Patterson, editors (1973). Useful host cell lines include COS-7, COS-A2, CV-1, murine cells such as murine myelomas N51 and VERO, HeLa cells, and Chinese hamster ovary (CHO) cells. Expression vectors for such cells ordinarily include promoters and control sequences  
25 compatible with mammalian cells such as, for example, the commonly used early and late promoters from Simian Virus 40 (SV 40), *see* Fiers *et al.*, 1978, *Nature* 273:113, or other viral promoters such as those derived from polyoma, adenovirus 2, bovine papilloma virus (BPV), or avian sarcoma viruses, or immunoglobulin promoters and heat shock promoters. A system for expressing DNA in mammalian systems using a BPV vector system is disclosed in United  
30 States Patent No. 4,419,446. A modification of this system is described in U.S. Patent No. 4,601,978. General aspects of mammalian cell host system transformations have been described by Axel, U.S. Patent No. 4,399,216. "Enhancer" regions are also important in optimizing expression; these are, generally, sequences found upstream of the promoter region.

Origins of replication may be obtained, if needed, from viral sources. However, integration into the chromosome is a common mechanism for DNA replication in eucaryotes.

Plant cells can also be used as hosts, and control sequences compatible with plant cells, such as the nopaline synthase promoter and polyadenylation signal sequences, *see* Depicker *et al.*, 1982, J. Mol. Appl. Gen. 1:561, are available. Expression systems employing insect cells utilizing the control systems provided by baculovirus vectors have also been described. *See* Miller *et al.*, in Genetic Engineering (1986), Setlow *et al.*, eds., Plenum Publishing, Vol. 8, pp. 277-97. Insect cell-based expression can be accomplished in *Spodoptera frugiperda*. These systems are also successful in producing recombinant enzymes.

Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described by Cohen, 1972, Proc. Natl. Acad. Sci. USA 69:2110 is used for procaryotes or other cells that contain substantial cell wall barriers. Infection with *Agrobacterium tumefaciens*, *see* Shaw *et al.*, 1983, Gene 23:315, is used for certain plant cells. For mammalian cells, the calcium phosphate precipitation method of Graham *et al.*, 1978, Virology 52:546 is preferred. Transformations into yeast are carried out according to the method of Van Solingen *et al.*, 1977, J. Bact. 130:946, and Hsiao *et al.*, 1979, Proc. Natl. Acad. Sci. USA 76:3829.

It may be desirable to modify the sequence of a DNA encoding a polypeptide comprising all or part of a reporter sequence, binding sequence or reporter fusion of the invention to provide, for example, a sequence more compatible with the codon usage of the host cell without modifying the amino acid sequence of the encoded protein. Such modifications to the initial 5-6 codons may improve expression efficiency. DNA sequences which have been modified to improve expression efficiency, but which encode the same amino acid sequence, are considered to be equivalent and encompassed by the present invention.

A variety of site-specific primer-directed mutagenesis methods are available and well-known in the art. *See, e.g.*, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1989, second edition, chapter 15.51, "Oligonucleotide-mediated mutagenesis," which is incorporated herein by reference. The polymerase chain reaction (PCR) can be used to perform site-specific mutagenesis. In another technique now standard in the art, a synthetic oligonucleotide encoding the desired mutation is used as a primer to direct synthesis of a complementary nucleic acid sequence contained in a single-stranded

vector, such as pBSM13+ derivatives, that serves as a template for construction of the extension product of the mutagenizing primer. The mutagenized DNA is transformed into a host bacterium, and cultures of the transformed bacteria are plated and identified. The identification of modified vectors may involve transfer of the DNA of selected transformants to a nitrocellulose filter or other membrane and the "lifts" hybridized with kinased synthetic mutagenic primer at a temperature that permits hybridization of an exact match to the modified sequence but prevents hybridization with the original unmutagenized strand. Transformants that contain DNA that hybridizes with the probe are then cultured (the sequence of the DNA is generally confirmed by sequence analysis) and serve as a reservoir of the modified DNA.

Once the polypeptide has been expressed in a recombinant host cell, purification of the polypeptide may be desired. A variety of purification procedures can be used.

For long-term stability, the purified polypeptide can be stored in a buffer that contains one or more non-ionic polymeric detergents. Such detergents are generally those that have a molecular weight in the range of approximately 100 to 250,00 preferably about 4,000 to 200,000 daltons and stabilize the enzyme at a pH of from about 3.5 to about 9.5, preferably from about 4 to 8.5. Examples of such detergents include those specified on pages 295-298 of McCutcheon's Emulsifiers & Detergents, North American edition (1983), published by the McCutcheon Division of MC Publishing Co., 175 Rock Road, Glen Rock, NJ (USA), the entire disclosure of which is incorporated herein by reference. Preferably, the detergents are selected from the group comprising ethoxylated fatty alcohol ethers and lauryl ethers, ethoxylated alkyl phenols, octylphenoxy polyethoxy ethanol compounds, modified oxyethylated and/or oxypropylated straight-chain alcohols, polyethylene glycol monooleate compounds, polysorbate compounds, and phenolic fatty alcohol ethers. More particularly preferred are Tween 20™, a polyoxyethylated (20) sorbitan monolaurate from ICI Americas Inc. (Wilmington, DE), and Iconol™ NP-40, an ethoxylated alkyl phenol (nonyl) from BASF Wyandotte Corp. (Parsippany, NJ).

The following series of examples are presented by way of illustration and not by way of limitation on the scope of the invention.

## 6. EXAMPLES

The present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

#### Example 1: SGN17 His Scan Method

This example demonstrates that a binding sequence can be modified to generate a binding sequence with a higher binding affinity and a binding sequence with a pH dependent binding affinity.

#### pADEPT06 DNA Template:

A schematic of plasmid pADEPT06 is shown in Figure 8. This plasmid is 5.2 kb and encodes a single chain antibody variable region fragment (scFv) fused to  $\beta$ -lactamase (BLA) with a pelB leader sequence, and is driven by a lac promoter (P lac). The plasmid also carries a chloramphenicol resistance gene (CAT) as a selectable marker. This particular SGN17 plasmid was made by a 3-piece ligation utilizing a linker. Two plasmids were used to make pADEPT06: pCB04 for the vector fragment with the pel B leader sequence, and pCR13 for the scFv-bla gene. pCB04 was digested with *HindIII* and *DraIII* (both from New England Biolabs, Beverly, MA) resulting in a 2.7kb fragment with the pCB04 backbone. pCR13 was digested with *NdeI* (Roche Molecular Biochemicals, Indianapolis, Indiana) and *DraIII* resulting in the 2.4kb fragment containing the fusion protein with the pelB leader sequence. Digests pCR13 were done in NEB2 buffer from NEB (50mM NaCl, 10mM Tris-HCl, 10mM MgCl<sub>2</sub>, 1mM dithiothreitol (pH 7.9 @ 25° C). Both fragments were gel purified from 1% agarose gel using a Qiagen kit (Qiagen, Valencia, CA). A linker sequence with 5' *HindIII* complementary ends and 3' *NdeI* complementary ends was used to link the 2.7kb fragment and the 2.4kb fragment upstream of the pel B leader sequence. The pCB04 fragment was combined with the pCR13 fragment and the linker in a 1:1:10 molar ratio (respectively), using 17  $\mu$ l DNA volume (95ng total DNA) and 17  $\mu$ l Takara ligase solution I (Panvera, Madison, WI) and incubated overnight at 16° C in a PTC-200™ machine (MJ Research, Waltham, MA). Sequencing information shows that the linker region is repeated upstream of the leader sequence.

#### Mutagenic Primers:

Overlapping mutagenic primers were designed to replace certain amino acids with histidine residues in the CDR3 regions of both the heavy and light chains of the scFv portion of the scFv-BLA fusion. The wild-type codon to be mutated was changed to the codon CAT (encoding histidine) in a pair of primers. The mutated codon in each primer was flanked on each side by 17 nucleotides of wild-type sequence, unless the primer ended in a stretch of A residues; in this case, the flanking sequence was extended so that it ended with a G or C residue. Each primer was designed so that its mutant codon had the same number of nucleotides flanking it on each side. Each primer was named according to the mutation it was designed to create. For example, HCL100F is the forward primer for the heavy chain (HC) mutating the Leucine (L) in position 100. Its overlapping primer is called HCL100R.

The names and sequences of the mutagenic oligos are provided in Table 1.

Table 1  
SGN17 His Scan Primers

15	Heavy Chain	
	HCK64F	ACTACAATCCATCTCTCC <u>CAT</u> AGTCGCAATTTCCATCAC
	HCK64R	GTGATGGAAATGCGACTATGGAGAGATGGATTGTAGT
20	HCR97F	GCCACATATTACTGTGCA <u>CAT</u> AGGACTCTGGCTACTTAC
	HCR97R	GTAAGTAGCCAGAGTCCTATGTGCACAGTAATATGTGGC
	HCR98F	CATATTACTGTGCAAGA <u>CAT</u> ACTCTGGCTACTTACTA
	HCR98R	TAGTAAGTAGCCAGAGTATGTCTTGACAGTAATATG
25	HCT99F	ATTACTGTGCAAGAAGG <u>CAT</u> CTGGCTACTTACTATGC
	HCT99R	GCATAGTAAGTAGCCAGATGCCTTCTTGACAGTAAT
	HCL100F	ACTGTGCAAGAAGGACT <u>CAT</u> GCTACTTACTATGCTAT
	HCL100R	ATAGCATAGTAAGTAGCATGAGTCCTTCTTGACAGT
30	HCA101F	GTGCAAGAAGGACTCTG <u>CAT</u> ACTTACTATGCTATGGA
	HCA101R	TCCATAGCATAGTAAGTATGCAGAGTCCTTCTTGAC
35	HCT102F	CAAGAAGGACTCTGGCT <u>CAT</u> TACTATGCTATGGACTA
	HCT102R	TAGTCCATAGCATAGTAATGAGCCAGAGTCCTTCTTG
	HCY103F	GAAGGACTCTGGCTACT <u>CAT</u> TATGCTATGGACTACTG
	HCY103R	CAGTAGTCCATAGCATAATGAGTAGCCAGAGTCCTTC
40	HCY104F	GGACTCTGGCTACTTACC <u>CAT</u> GCTATGGACTACTGGGG
	HCY104R	CCCCAGTAGTCCATAGCATGGTAAGTAGCCAGAGTCC
	HCA105F	CTCTGGCTACTTACTAT <u>CAT</u> ATGGACTACTGGGGTCA
	HCA105R	TGACCCAGTAGTCCATATGATAGTAAGTAGCCAGAG
45	HCM106F	TGGCTACTTACTATGCT <u>CAT</u> GACTACTGGGGTCAAGG
	HCM106R	CCTTGACCCAGTAGTCATGAGCATAGTAAGTAGCCA
50	HCD107F	CTACTTACTATGCTATG <u>CAT</u> TACTGGGGTCAAGGAAC
	HCD107R	GTTCCTTGACCCAGTAATGCATAGCATAGTAAGTAG

	HCY108F	CTTACTATGCTATGGAC <b><u>CAT</u></b> TGGGGTCAAGGAACCTC
	HCY108R	GAGGTTCCCTTGACCCCAATGGTCCATAGCATAGTAAG
5	HCW109F	ACTATGCTATGGACTAC <b><u>CAT</u></b> TGGTCAAGGAACCTCTGT
	HCW109R	ACAGAGGTTCCCTTGACCATGGTAGTCCATAGCATAGT
	Light Chain	
	LCR54F	CAAAGCTCCTGATCTAC <b><u>CAT</u></b> GTTTCCAACCGATTTTC
10	LCR54R	GAAAATCGGTTGGAAACATGGTAGATCAGGAGCTTTG
	LCR58F	GATTTTCTGGGGTCCCAGAC <b><u>CAT</u></b> TTTCAGTGGCAGTGGATCAGG
	LCR58R	CCTGATCCACTGCCACTGAAATGGTCTGGGACCCAGAAAATC
15	LCQ94F	GAGTTTATTTCTGCTCT <b><u>CAT</u></b> AGTACACATGTTCTCTCC
	LCQ94R	GGAGGAACATGTGTACTATGAGAGCAGAAATAAACTC
	LCS95F	GTTTATTTCTGCTCTCAA <b><u>CAT</u></b> ACACATGTTCTCTCCGACG
	LCS95R	CGTCGGAGGAACATGTGTATGTTGAGAGCAGAAATAAAC
20	LCT96F	GTTTATTTCTGCTCTCAAAGT <b><u>CAT</u></b> CATGTTCTCTCCGACGTTCCGGT
	LCT96R	ACCGAACGTCGGAGGAACATGATGACTTTGAGAGCAGAAATAAAC
	LCH97F	TCTGCTCTCAAAGTACAC <b><u>CAT</u></b> GTTCTCTCCGACGTTCCGG
25	LCH97R	CCGAACGTCGGAGGAACATGTGTACTTTGAGAGCAGA
	LCV98F	GCTCTCAAAGTACACAT <b><u>CAT</u></b> CTCTCCGACGTTCCGGTGG
	LCV98R	CCACCGAACGTCGGAGGATGATGTGTACTTTGAGAGC
30	LCP99F	CTCAAAGTACACATGTT <b><u>CAT</u></b> CCCGACGTTCCGGTGGAGG
	LCP99R	CCTCCACCGAACGTCGGATGAACATGTGTACTTTGAG
	LCP100F	CAAAGTACACATGTTCTCT <b><u>CAT</u></b> ACGTTCCGGTGGAGGCACC
	LCP100R	GGTGCCTCCACCGAACGTATGAGGAACATGTGTACTTTG
35	LCT101F	AGTACACATGTTCTCTCCG <b><u>CAT</u></b> TTCCGGTGGAGGCACCAAG
	LCT101R	CTTGGTGCCTCCACCGAAATGCCGAGGAACATGTGTACT

All sequences written 5'-3'. Mutagenic codon in bold and underlined.

40 A QUICKCHANGE™ site directed mutagenesis kit (Stratagene, La Jolla, CA)  
was used to set up PCR amplifications as follows:

	H <sub>2</sub> O	39 µl
	10x buffer	5 µl
45	dNTP mix	1.5 µl
	Forward primer	1 µl (0.5 µM final concentration)
	Reverse primer	1 µl (0.5 µM final concentration)
	<i>pfu</i> polymerase	1 µl
	Plasmid DNA	1.5 µl (150 ng)
50	Total	50 µl

The buffer comprised 100mM KCl, 100mM (NH<sub>4</sub>)SO<sub>4</sub>, 200mM Tris-HCl (pH 8.8), 20mM MgSO<sub>4</sub>, 1% Triton X-100, 1mg/ml nuclease-free bovine serum albumin (BSA).

The following touchdown PCR program was used in a PTC-200™ machine (MJ Research, Waltham, MA):

- 5                   1)     95° C, 2 minutes
- 2)     95° C, 45 seconds
- 3)     60° C, 1 minute  
                          (Reduced by 1.0 °C per cycle)
- 10               4)     68° C 11 minutes (*i.e.*, 2 minutes per kb, 5 kb plasmid, plus an  
                          additional minute)
- 5)     Go to step (2) for 9 cycles
- 6)     95° C, 45 seconds
- 7)     50° C, 1 minute
- 8)     68° C, 11 minutes
- 15               9)     Go to step (6) for 5 cycles
- 10)    Hold at 4° C

A negative control without primers was also set up and carried through all steps.

20

#### *DpnI* digest:

*DpnI* is a restriction enzyme that cuts methylated and hemimethylated, but not unmethylated, double-stranded DNA. After PCR, 1 µl of *DpnI* was added to each reaction to digest template DNA, which is methylated, but not amplified DNA, most of which is  
25   unmethylated, thus reducing the background of wild-type sequence. A sample of the control was saved before digestion. Digests were incubated at 37° C for 1.5 hrs, then each reaction was spiked with an additional 1 µl of *DpnI* and incubated another 1.5 hrs. Reactions were run on a gel after digests alongside the control amplification before and after *DpnI* digestion. All reactions appeared to work, and, as expected, the control band was fully digested by *DpnI*.

30

#### Transformation:

1 µl of each reaction (not purified), including the digested control, were used to transform 50 µl of Top 10 electro-competent cells (Invitrogen, Carlsbad, CA) and 250 µl SOC medium (2% Bacto-Tryptone, 0.5% Bacto Yeast Extract, 10 mM NaCl, 2.5 mM KCl) was  
35   added. The cells were shaken at 37° C for 45 min, then 30 µl of a 1 to 10 dilution was plated (*i.e.*, one tenth of the total volume of each transformation was plated) on both 5ppm chloramphenicol (CMP) and 5ppm CMP + 0.1ppm cefotaxime (CTX) plates. Plates were incubated overnight at 37° C. Transformation results are provided in Table 2.



		Table 2		
		CMP	CMP + CTX	% ACTIVE
	(control)	0	0	0
	ME43	14	5	36
5	ME44	120	34	28
	ME45	784	236	32
	ME46	440	159	36
	ME47	516	184	36
	ME48	268	62	23
10	ME49	30	10	33
	ME50	488	61	12.5
	ME51	316	57	18
	ME52	380	192	50
	ME53	440	80	18
15	ME54	968	308	32
	ME55	356	148	42
	ME56	90	17	19
	ME57	424	112	26
	ME58	38	10	26
20	ME59	141	53	38
	ME60	212	144	68
	ME61	90	27	30
	ME62	268	87	32 (WT codon)
	ME63	296	88	30
25	ME64	196	112	57
	ME65	168	128	76
	ME66	236	76	32

All bacteria transformed by and expressing a plasmid produced colonies on the CTX plate, and thus provided a measure of the efficiency of transformation. However, only bacteria transformed by plasmids containing a functional BLA grew on the CTX + CMP plates.

Clone names in Table 2 are listed in the same order as the primer pairs used to make them are listed in Table 1, *e.g.*, ME43 was created using primer pair HCK64F/R, ME44 was created using primer pair HCR97F/R, and so on.

Four colonies were picked for each transformation (excluding LCH97 because this represents the wild-type sequence; pADEPT06 WT colonies were picked as a control). Picked colonies were first swirled into a 96 well plate with membrane bottom, each well containing 200ul LB + 5ppm CMP, and then put into the corresponding well of another 96 well plate without filter, to be used as a stock plate.

The 96 well plates were incubated at 25° C in a humidified box with shaking for 48 hrs. Glycerol was added to the stock plate to a final concentration of 10% and stored at -80° C.

Screening mutants:

Target protein p97 (prepared, for example, by the method set forth in Siemers, N.O., D.E. Kerr, S. Yarnold, M.R. Stebbins, V.M. Vrudhula, I. Hellstrom and P.D. Senter (1997), *Bioconj Chem* 8, 510-519., Construction, expression and activities of L49-sFv-beta-lactamase, a single-chain antibody fusion protein for anticancer prodrug activation) was

5 immobilized on a polystyrene plate by adding 100  $\mu$ l of 1  $\mu$ g/ml p97 in PBS and incubating the plate at 4° C overnight. The plate is then washed with PBST (PBS+0.25%Tween 20) and blocked with 200  $\mu$ l/well of 1% casein in PBS overnight at 4° C. On the day of screening, the plate was washed with PBST, then each well received 80  $\mu$ l of 50mM PBS pH7.4 and 20  $\mu$ l of cell culture broth from each mutant. The plate was incubated at room temperature with gentle  
10 shaking to let SGN-17 bind to immobilized p97 on the plate. The amount of each mutant enzyme bound to p97 was determined at two time points. After 1 hour, the plate was washed with PBST, and 200  $\mu$ l of the BLA substrate nitrocefin in 50mM PBS buffer pH7.4 or pH6.5 was added into each well. The amount of bound SGN-17 was measured by monitoring hydrolysis of nitrocefin at wavelength 490nm. This was the T<sub>0</sub> time point measurement. The  
15 plate was then incubated in each substrate buffer for one hour, providing an opportunity for bound mutant SGN-17 to dissociate, then quickly rinsed with PBST. The remaining bound SGN-17 was measured by again monitoring the hydrolysis of substrate nitrocefin in each buffer. This was the T<sub>1</sub> time point measurement. A ratio of bound activity at T<sub>1</sub> vs. T<sub>0</sub> was calculated for each mutant, and an index was calculated by dividing the ratio of mutant over  
20 parent, as shown in Table 3.

Table 3

	Mutants	sequence	position	region	Index pH7.4	Index pH6.5
	ME43	K	HC62	CDR2	0.61	0.65
25	ME44	R	HC94	CDR3	0.24	0
	ME45	R	HC95	CDR3	0	0
	ME46	T	HC96	CDR3	0.38	0.09
	ME47	L	HC97	CDR3	0.24	0
	ME48	A	HC98	CDR3	0.49	0.33
30	ME50	Y	HC100	CDR3	0.33	0
	ME51	Y	HC101	CDR3	0.26	0
	ME52	A	HC102	CDR3	0	0
	ME53	M	HC103	CDR3	0.97	0.8
	ME54	D	HC104	CDR3	0.41	0.7
35	ME55	Y	HC105	CDR3	0.8	0.7
	ME56	W	HC106	CDR3	0.57	0.41
	ME58	R	LC58	CDR2	0.92	0.76
	ME59	Q	LC94	CDR3	0.28	0
	ME60	S	LC95	CDR3	1.04	1.09
40	ME61	T	LC96	CDR3	0.82	0.81
	ME63	V	LC98	CDR3	0.21	0

ME64	P	LC99	CDR3	0.35	0
ME65	P	LC100	CDR3	0	0
ME66	T	LC101	CDR3	1.36	1.73

5 A high index value for a mutant indicates that it has a slow  $k_{off}$ . An index value of 0 indicates that no binding was detected for the mutant at that pH.

10 These data illustrate that many residues in the CDR3 of SGN-17 can be replaced with His while retaining various degrees of binding affinity. Mutagenesis at position LC101 actually leads to an increase in binding affinity which is larger at pH 6.5 as compared to pH 7.4. Thus, the introduction of a His in position LC101 affects the pH-dependence of target binding of SGN-17. Comparing the index values at both pH values shows that several of the tested mutations affect pH-dependence of binding. Stronger effects can be achieved by adding further mutations, by testing substitutions other than His, by testing substitutions, insertions or deletions at more positions of the binding moiety, or by extending the  
15 mutagenesis to the BLA part of the fusion protein.

#### Example 2: Affinity maturation of an scFv by site saturation scanning mutagenesis

##### A. Generation of site saturation libraries

64 site saturation mutagenesis libraries were generated. In each of these libraries, one codon, that codes for a CDR position (as defined by the Kabat nomenclature) in ME66.4-scFv, exactly the same as ME66, was randomized. The libraries were generated using the QuikChange protocol (Stratagene, La Jolla, CA) essentially as recommended by the manufacturer. Each reaction used two mutagenic oligonucleotides which had the following design: 17 perfect matches flanking the random codon on each side, NNS in place of the  
25 random codon. For example, library ME67 used the forward primer CTGGCGACTCCATCACCNNSGGT TACTGGA ACTGGAT and the reverse primer ATCCAGTTCCAGTAACCSNNGGTGATGGAGTCGCCAG, where N represents a mixture of A, T, G, and C and S represents a mixture of G and C. This approach allows for the generation of 32 different codons which encode all 20 amino acids. After the QuikChange  
30 reaction and Dpn I digest, which degrades parent plasmid, the reaction mixture was used to transform TOP10 cells (Invitrogen, Carlsbad, CA) by electroporation.

Table 4: oligonucleotides used to generate the 64 site saturation libraries:

##### Heavy Chain

H31 CTGGCGACTCCATCACCNNSGGT TACTGGA ACTGGAT

H31 ATCCAGTTCAGTAACCSNNGGTGATGGAGTCGCCAG  
H32 GCGACTCCATCACCAGTNNSTACTGGAAGTGGATCCG  
H32 CGGATCCAGTTCAGTASNNACTGGTATGGAGTCGC  
H33 ACTCCATCACCAGTGGTNNSTGGAAGTGGATCCGGCA  
H33 TGCCGGATCCAGTTCASNNACCACTGGTATGGAGT  
H34 TCCATCACCAGTGGTTACNNSAAGTGGATCCGGCAGTTC  
H34 GAACTGCCGGATCCAGTTSNNGTAACCACTGGTATGGA  
H50 AACTTGAATATATGGGTNNSATAAGCGACAGTGGTAT  
H50 ATACCACTGTCGCTTATSNNACCCATATATTCAAGTT  
H51 TTGAATATATGGGTTACNNSAGCGACAGTGGTATCAC  
H51 GTGATACCACTGTCGCTSNNGTAACCCATATATTCAA  
H52 GAATATATGGGTTACATANNSGACAGTGGTATCACTTAC  
H52 GTAAGTGATACCACTGTCNNATATGTAACCCATATATTC  
H53 TATATGGGTTACATAAGCNNSAGTGGTATCACTTACTAC  
H53 GTAGTAAGTGATACCACTSNNGCTTATGTAACCCATATA  
H54 ATGGGTTACATAAGCGACNNSGGTATCACTTACTACAAT  
H54 ATTGTAGTAAGTGATACCSNNGTCGCTTATGTAACCCAT  
H55 GTTACATAAGCGACAGTNNSATCACTTACTACAATCC  
H55 GGATTGTAGTAAGTGATSNNACTGTCGCTTATGTAAC  
H56 ACATAAGCGACAGTGGTNNSACTTACTACAATCCATC  
H56 GATGGATTGTAGTAAGTSNNACCACTGTCGCTTATGT  
H57 TAAGCGACAGTGGTATCNNSTACTACAATCCATCTCT  
H57 AGAGATGGATTGTAGTASNNGATACCACTGTCGCTTA  
H58 TAAGCGACAGTGGTATCACTNNSTACAATCCATCTCTCAAAAG  
H58 CTTTTGAGAGATGGATTGTASNAGTGATACCACTGTCGCTTA  
H59 GACAGTGGTATCACTTACNNSAATCCATCTCTCAAAAGT  
H59 ACTTTTGAGAGATGGATTSNNGTAAGTGATACCACTGTC  
H60 GTGGTATCACTTACTACNNSCCATCTCTCAAAAGTCG  
H60 CGACTTTTGAGAGATGGSNNGTAGTAAGTGATACCAC  
H61 GTATCACTTACTACAATNNSTCTCTCAAAAGTCGCAT  
H61 ATGCGACTTTTGAGAGASNATTGTAGTAAGTGATAC  
H62 TCACTTACTACAATCCANNSTCAAAAGTCGCATTTC  
H62 GAAATGCGACTTTTGAGSNNTGGATTGTAGTAAGTGA  
H63 CTTACTACAATCCATCTNNSAAAAGTCGCATTTCCAT  
H63 ATGGAAATGCGACTTTTSSNAGATGGATTGTAGTAAG  
H64 ACTACAATCCATCTCTCNNSAGTCGCATTTCCATCAC  
H64 GTGATGGAAATGCGACTSNNGAGAGATGGATTGTAGT  
H65 ACAATCCATCTCTCAAANNSTCGCATTTCCATCACTCG  
H65 CGAGTGATGGAAATGCGSNNTTTGAGAGATGGATTGT  
H97 GCCACATATTACTGTGCANNSAGGACTCTGGCTACTTAC

H97 GTAAGTAGCCAGAGTCCTSNNTGCACAGTAATATGTGGC  
 H98 CATATTACTGTGCAAGANNSACTCTGGCTACTTACTA  
 H98 TAGTAAGTAGCCAGAGTSNNTCTTGCACAGTAATATG  
 H99 ATTACTGTGCAAGAAGGNNNSCTGGCTACTTACTATGC  
 H99 GCATAGTAAGTAGCCAGSNNCCTTCTTGCACAGTAAT  
 H100 ACTGTGCAAGAAGGACTNNSGCTACTTACTATGCTAT  
 H100 ATAGCATAGTAAGTAGCSNNAGTCCTTCTTGCACAGT  
 H101 GTGCAAGAAGGACTCTGNNSACTTACTATGCTATGGA  
 H101 TCCATAGCATAGTAAGTSNNCAGAGTCCTTCTTGCAC  
 H102 CAAGAAGGACTCTGGCTNNSTACTATGCTATGGACTA  
 H102 TAGTCCATAGCATAGTASNNAGCCAGAGTCCTTCTTG  
 H103 GAAGGACTCTGGCTACTNNSTATGCTATGGACTACTG  
 H103 CAGTAGTCCATAGCATASNNAGTAGCCAGAGTCCTTC  
 H104 GGACTCTGGCTACTTACNNSGCTATGGACTACTGGGG  
 H104 CCCCAGTAGTCCATAGCSNNGTAAGTAGCCAGAGTCC  
 H105 CTCTGGCTACTTACTATNNSATGGACTACTGGGGTCA  
 H105 TGACCCCAGTAGTCCATSNNATAGTAAGTAGCCAGAG  
 H106 TGGCTACTTACTATGCTNNSGACTACTGGGGTCAAGG  
 H106 CCTTGACCCCAGTAGTCSNNAGCATAGTAAGTAGCCA  
 H107 CTACTTACTATGCTATGNNSTACTGGGGTCAAGGAAC  
 H107 GTTCCTTGACCCCAGTASNNCATAGCATAGTAAGTAG  
 H108 CTTACTATGCTATGGACNNSTGGGGTCAAGGAACCTC  
 H108 GAGGTTCTTGACCCASNNGTCCATAGCATAGTAAG  
 H109 ACTATGCTATGGACTACNNSGGTCAAGGAACCTCTGT  
 H109 ACAGAGGTTCTTGACCSNNGTAGTCCATAGCATAGT

**Light Chain**

L24 CCTCCATCTCTTGCAGGNNNSAGTCAGAGCCTTGTACA  
 L24 TGTACAAGGCTCTGACTSNNCCTGCAAGAGATGGAGG  
 L25 CCATCTCTTGCAGGGCTNNSCAGAGCCTTGTACACAG  
 L25 CTGTGTACAAGGCTCTGSNNAGCCCTGCAAGAGATGG  
 L26 ATCTCTTGCAGGGCTAGTNNNSAGCCTTGTACACAGTAAT  
 L26 ATTACTGTGTACAAGGCTSNNACTAGCCCTGCAAGAGAT  
 L27 CTTGCAGGGCTAGTCAGNNSCTTGTACACAGTAATGG  
 L27 CCATTACTGTGTACAAGSNNCTGACTAGCCCTGCAAG  
 L28 TGCAGGGCTAGTCAGAGCNSGTACACAGTAATGGAAAC  
 L28 GTTTCCATTACTGTGTACSNNGCTCTGACTAGCCCTGCA  
 L29 GGGCTAGTCAGAGCCTTNNSCAGTAATGGAAACAC  
 L29 GTGTTTCCATTACTGTGSNNAAGGCTCTGACTAGCCC  
 L30 CTAGTCAGAGCCTTGTANNSAGTAATGGAAACACCTA  
 L30 TAGGTGTTTCCATTACTSNNTACAAGGCTCTGACTAG

L31 TAGTCAGAGCCTTGTACACNNSAATGGAAACACCTATTTAC  
L31 GTAAATAGGTGTTTCCATTSNNGTGTACAAGGCTCTGACTA  
L32 AGAGCCTTGTACACAGTNNSGGAAACACCTATTTACA  
L32 TGTAATAGGTGTTTCCSNNACTGTGTACAAGGCTCT  
L33 GCCTTGTACACAGTAATNNSAACACCTATTTACATTG  
L33 CAATGTAAATAGGTGTTSNNATTACTGTGTACAAGGC  
L34 TTGTACACAGTAATGGANNSACCTATTTACATTGGTA  
L34 TACCAATGTAAATAGGTSNNTCCATTACTGTGTACAA  
L35 TACACAGTAATGGAAACNNSTATTTACATTGGTACC  
L35 GGTACCAATGTAAATASNNGTTTCCATTACTGTGTA  
L36 ACAGTAATGGAAACACCNNSTTACATTGGTACCTGCA  
L36 TGCAGGTACCAATGTAASNNGGTGTTTCCATTACTGT  
L37 AGTAATGGAAACACCTATNNSCATTGGTACCTGCAGAAG  
L37 CTTCTGCAGGTACCAATGSNNATAGGTGTTTCCATTACT  
L38 ATGGAAACACCTATTTANNSTGGTACCTGCAGAAGCC  
L38 GGCTTCTGCAGGTACCASNNTAAATAGGTGTTTCCAT  
L53 CTCCAAAGCTCCTGATCNSAGAGTTTCCAACCGATT  
L53 AATCGGTTGGAACTCTSNNGATCAGGAGCTTTGGAG  
L54 CAAAGCTCCTGATCTACNNSGTTTCCAACCGATTTTC  
L54 GAAAATCGGTTGGAAACSNNGTAGATCAGGAGCTTTG  
L55 AGCTCCTGATCTACAGANNSTCCAACCGATTTTCTGG  
L55 CCAGAAAATCGGTTGGASNNTCTGTAGATCAGGAGCT  
L56 TCCTGATCTACAGAGTTNNSAACCGATTTTCTGGGGT  
L56 ACCCCAGAAAATCGGTTSNNACTCTGTAGATCAGGA  
L57 TGATCTACAGAGTTTCCNNSCGATTTTCTGGGGTCCC  
L57 GGGACCCAGAAAATCGSNNGGAACTCTGTAGATCA  
L58 TCTACAGAGTTTCCAACNNSTTTTCTGGGGTCCCAGA  
L58 TCTGGGACCCAGAAAASNNGTTGGAACTCTGTAGA  
L59 ACAGAGTTTCCAACCGANNSTCTGGGGTCCCAGACAG  
L59 CTGTCTGGGACCCAGASNNTCGGTTGGAACTCTGT  
L60 GAGTTTCCAACCGATTTNNSGGGGTCCCAGACAGGTT  
L60 AACCTGTCTGGGACCCCSNNAATCGGTTGGAACTC  
L94 GAGTTTATTTCTGCTCTNNSAGTACACATGTTCTCTCC  
L94 GGAGGAACATGTGTACTSNNAGAGCAGAAATAAACTC  
L95 GAGTTTATTTCTGCTCTCAANNSACACATGTTCTCTCCGATT  
L95 AAATGCGGAGGAACATGTGTSNNTTGAGAGCAGAAATAAACTC  
L96 TATTTCTGCTCTCAAAGTNNSCATGTTCTCTCCGATTTCC  
L96 GAAATGCGGAGGAACATGSNNACTTTGAGAGCAGAAATA  
L97 TCTGCTCTCAAAGTACANNSGTTCTCTCCGATTTCCG  
L97 CCGAAATGCGGAGGAACSNNTGTACTTTGAGAGCAGA

L98 GCTCTCAAAGTACACATNNSCCTCCGCATTTCCGGTGG  
 L98 CCACCGAAATGCGGAGGSNNATGTGTACTTTGAGAGC  
 L99 CTCAAAGTACACATGTTNNSCCGCATTTCCGGTGGAGG  
 L99 CCTCCACCGAAATGCGGSNNAACATGTGTACTTTGAG  
 L100 CAAAGTACACATGTTCTNNSCATTTCGGTGGAGGCACC  
 L100 GGTGCCTCCACCGAAATGSNNAGGAACATGTGTACTTTG  
 L101 AGTACACATGTTCTCCGNNSTTCGGTGGAGGCACCAAG  
 L101 CTTGGTGCCTCCACCGAASNCGGAGGAACATGTGTACT

#### B. Screen for improved binding

Libraries were plated onto agar plates containing LB medium and 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime (Sigma). 88 colonies from each library and parent colonies were picked and inoculated into 384-well plates containing 80 ul LB containing 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime. Plates were incubated at 25 C in humidified boxes with shaking for 48 hrs.

Target protein p97 was immobilized in 384-well polystyrene plates by adding 40 ul of 1 ug/ml p97 in PBS and incubating the plate at 4C overnight. The plates were then washed with PBST (PBS+0.1%Tween-20) and blocked with 200ul/well of 1% Casein in PBS overnight at 4C. On the day of screening, the plates were washed with PBST. Subsequently, 24ul/well of 50mM PBS pH7.4 was first added into plate each well followed by 8 ul of cell culture broth from expression plates. The plate was incubated at room temperature with gentle shaking to let ME66-scFv to bind to immobilized P97 on the plate. After 1 hour, the plate was washed with PBST and 200ul of BLA assay buffer containing 0.1 mg/ml nitrocefin (Oxoid, New York) in 50mM PBS buffer pH6.5 was added into each well, the bound ME66scFv was measured by monitoring hydrolysis of nitrocefin at wavelength 490nm. The plate was then left incubated in substrate buffer to allow the bound ME66scFv-BLA to dissociate, after 1.5 hour the plate was quickly rinsed with PBST. The remaining bound ME66scFv-BLA was again measured by monitoring the hydrolysis of freshly added substrate nitrocefin. Dissociation of ME66-scFv from p97 was monitored again after 3-5 hours. A ratio of bound activity at time 1 vs. time 0 or time 2 vs. time 0 was calculated for each mutant from dissociation data, an index at each time point was further calculated by dividing ratio of mutant over parent, and winner mutants were chosen if they had a high index.

After the primary screening, 21 winners were chosen for repeat analysis in quadruplicates. Each winner was streaked out on LA agar containing 5 mg/l chloramphenicol, 4 colonies from each winner were transferred in 96 well plate containing 200ul/well LB containing 5 mg/l

chloramphenicol. Some wells were inoculated with ME66.4 as a reference. The plate was incubated at 25C for 70 hours. Target protein p97 was biotinylated and immobilized in 96 well neutravidin (Pierce, Rockford, Il) plate at a p97 concentration of 5ug/ml of 100ul/well, the plate was then blocked with 1% Casein. On the day of screening, 70ul/well of PBS buffer pH7.4 was added into target plate, and 20ul/well of culture broth was transferred from expression plate to target plate. The plates were incubated at room temperature for 1 hour, and were then washed with PBST. 200ul of BLA substrate nitrocefin in 50mM PBS buffer pH6.5 were added into each well, and the bound ME66scFv was measured by monitoring hydrolysis of nitrocefin at wavelength 490nm. The plate was left incubated in substrate buffer for an additional 1.5 hour. After quick rinsing with PBST, the bound ME66scFv-BLA was again measured using nitrocefin. The dissociation of ME66scFv from p97 was again measured between 3-6 hours after the initial time point and a binding index was calculated at 2 time points. In parallel, the plate was screened under identical conditions but using 50 mM PBS buffer at pH 7.4. Data were normalized as described in Example 1. The normalized screening results measured at pH 6.5 and at pH 7.4 are shown in the Figure 9.

Table 6, below, shows mutations that have been observed in the three best variants.

Table 4: Mutations in affinity matured variants

Clone	mutation
ME70.1	heavy chain, S65K
ME70.7	heavy chain, S65P
ME81.3	heavy chain, N60R

### Example 3: Stabilization of an scFv

#### A. Construction of pME27.1



Plasmid pME27.1 was generated by inserting a Bgl I EcoRV fragment encoding a part of the pelB leader, the CAB1-scFv and a small part of BLA into plasmid the expression vector pME25. The insert, encoding for the CAB1-scFv, has been synthesized by Aptagen (Herndon, VA) based on the sequence of the scFv MFE-23 that was described in [Boehm, M. K., A. L. Corper, T. Wan, M. K. Sohi, B. J. Sutton, J. D. Thornton, P. A. Keep, K. A. Chester, R. H. Begent and S. J. Perkins (2000) *Biochem J* 346 Pt 2, 519-28, Crystal structure of the anti-(carcinoembryonic antigen) single-chain Fv antibody MFE-23 and a model for antigen binding based on intermolecular contacts]. Both the plasmid containing the synthetic gene (pPCR-GME1) and pME25 were digested with BglI and EcoRV, gel purified and ligated together with Takara ligase. Ligation was transformed into TOP10(Invitrogen, Carlsbad, CA) electrocompetent cells, plated on LA medium containing 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime.

Plasmid pME27.1 contains the following features:

**P lac:** 4992-5113 bp

**pel B leader:** 13-78

**CAB 1 scFv:** 79-810

**BLA:** 811-1896

**T7 term.:** 2076-2122

**CAT:** 3253-3912

A schematic of plasmid pME27.1 can be found in Figure 10A. The CAB1 sequence, indicating heavy and light chain domains, can be found in Figure 10B; the amino acid sequence can also be found in 10D, with linker and BLA.

#### B. Choosing mutations for mutagenesis

The sequence of the vH and vL sequences of CAB1-scFv were compared with a published frequency analysis of human antibodies (Boris Steipe (1998) Sequenzdatenanalyse. ("Sequence Data Analysis", available in German only) in *Bioanalytik* eds. H. Zorbas und F. Lottspeich, Spektrum Akademischer Verlag. S. 233-241). The authors aligned sequences of variable segments of human antibodies as found in the Kabat data base and calculated the

frequency of occurrence of each amino acid for each position. These alignment can be seen in Figure 12. Specifically, Figure 12A shows an alignment of the observed frequencies of the five most abundant amino acids in alignment of human sequences in the heavy chain. Figure 12B shows an alignment of the observed frequencies of the five most abundant amino acids in alignment of human sequences in the light chain.

We compared these frequencies with the actual amino acid sequence of CAB1 and identified 33 positions that fulfilled the following criteria:

- The position is not part of a CDR as defined by the Kabat nomenclature.
- The amino acid found in CAB1-scFv is observed in the homologous position in less than 10% of human antibodies
- The position is not one of the last 6 amino acids in the light chain of scFv.

The resulting 33 positions were chosen for combinatorial mutagenesis. Mutagenic oligonucleotides were synthesized for each of the 33 positions such that the targeted position would be changed from the amino acid in CAB1-scFv to the most abundant amino acid in the homologous position of a human antibody. Figure 10B shows the sequence of CAB1-scFv, the CDRs, and the mutations that were chosen for combinatorial mutagenesis.

### C. Construction of library NA05

Table 5 listing the sequences of 33 mutagenic oligonucleotides that were used to generate combinatorial library NA05:

Table 5:

pos. (pME27)	MFE-23	(VH)	count residues to be changed	QuikChange multi primer
3K	Q	nsa147.1fp		CGGCCATGGCCCAGGTGCAGCTGCAGCAGTCTGGGGC
13R	K	nsa147.2fp		CTGGGGCAGAACTTGTGAAATCAGGGACCTCAGTCAA
14S	P	nsa147.3fp		GGGCAGAACTTGTGAGGCCGGGGACCTCAGTCAAGTT
16T	G	nsa147.4fp		AACTTGTGAGGTCAGGGGGCTCAGTCAAGTTGTCCTG
28N	T	nsa147.5fp		GCACAGCTTCTGGCTTCACCATTAAGACTCCTATAT
29I	F	nsa147.6fp		CAGCTTCTGGCTTCAACTTTAAAGACTCCTATATGCA
30K	S	nsa147.7fp		CTTCTGGCTTCAACATTAGCGACTCCTATATGCACTG
37L	V	nsa147.8fp		ACTCCTATATGCACTGGGTGAGGCAGGGGCCTGAACA

40G A nsa147.9fp TGCACTGGTTGAGGCAGGCGCCTGAACAGGGCCTGGA  
 42E G nsa147.10fp GGTGAGGCAGGGGCCTGGCCAGGGCCTGGAGTGGAT  
 67K R nsa147.11fp CCCC GAAGTTCCAGGGCCGTGCCACTTTTACTACAGA  
 68A F nsa147.12fp CGAAGTTCCAGGGCAAGTTCACTTTTACTACAGACAC  
 70F I nsa147.13fp TCCAGGGCAAGGCCACTATTACTACAGACACATCCTC  
 72T R nsa147.14fp GCAAGGCCACTTTTACTCGCGACACATCCTCCAACAC  
 76S K nsa147.15fp TTACTACAGACACATCCAAAAACACAGCCTACCTGCA  
 97N A nsa147.16fp CTGCCGTCTATTATTGTGCGGAGGGGACTCCGACTGG  
 98E R nsa147.17fp CCGTCTATTATTGTAATCGCGGGACTCCGACTGGGCC  
 136E Q nsa147.18fp CTGGCGGTGGCGGATCACAGAATGTGCTCACCCAGTC  
 137N S nsa147.19fp GCGGTGGCGGATCAGAAAGCGTGCTCACCCAGTCTCC  
 142S P nsa147.20fp GAAAATGTGCTCACCCAGCCGCCAGCAATCATGTCTGC  
 144A S nsa147.21fp TGCTCACCCAGTCTCCAAGCATCATGTCTGCATCTCC  
 146M V nsa147.22fp CCCAGTCTCCAGCAATCGTGTCTGCATCTCCAGGGGA  
 152E Q nsa147.23fp TGTCTGCATCTCCAGGGCAGAAAGGTCACCATAACCTG  
 153K T nsa147.24fp CTGCATCTCCAGGGGAGACCGTCACCATAACCTGCAG  
 170F Y nsa147.25fp TAAGTTACATGCACTGGTACCAGCAGAAGCCAGGCAC  
 181W V nsa147.26fp GCACTTCTCCCAAACCTCGTGATTTATAGCACATCCAA  
 194A D nsa147.27fp TGGCTTCTGGAGTCCCTGATCGCTTCAGTGGCAGTGG  
 200G K nsa147.28fp CTCGCTTCAGTGGCAGTAAATCTGGGACCTCTTACTC  
 205Y A nsa147.29fp GTGGATCTGGGACCTCTGCGTCTCTCACAATCAGCCG  
 212M L nsa147.30fp CTCTCACAATCAGCCGACTGGAGGCTGAAGATGCTGC  
 217A E nsa147.31fp GAATGGAGGCTGAAGATGAAGCCACTTATTACTGCCA  
 219T D nsa147.32fp AGGCTGAAGATGCTGCCGATTATTACTGCCAGCAAAG  
 234A G nsa147.33fp ACCCACTCACGTTCGGTGGCGGCACCAAGCTGGAGCT

The QuikChange multi site-directed mutagenesis kit (QCMS; Stratagene Catalog # 200514) was used to construct the combinatorial library NA05 using 33 mutagenic primers. The primers were designed so that they had 17 bases flanking each side of the codon of interest based on the template plasmid pME27.1. The codon of interest was changed to encode the appropriate consensus amino acid using an *E.coli* codon usage table. All primers were designed to anneal to the same strand of the template DNA (i.e., all were forward primers in this case). The QCMS reaction was carried out as described in the QCMS manual with the

exception of the primer concentration used, which recommends using 50ng of each primer in the reaction whereas we used around 3 ng of each primer. Other primer amounts may be used. In particular, the reaction contained 50-100 ng template plasmid (pME27.1; 5178bp), 1 µl of primers mix (10 µM stock of all primers combined containing 0.3 µM each primer), 1 µl dNTPs (QCMS kit), 2.5 µl 10x QCMS reaction buffer, 18.5 µl deionized water, and 1 µl enzyme blend (QCMS kit), for a total volume of 25 µl. The thermocycling program was 1 cycle at 95° for 1 min., followed by 30 cycles of 95°C for 1 min., 55°C for 1 min., and 65°C for 10 minutes. *DpnI* digestion was performed by adding 1 µl *DpnI* (provided in the QCMS kit), incubation at 37°C for 2 hours, addition of another 1 µl *DpnI*, and incubation at 37°C for an additional 2 hours. 1 µl of the reaction was transformed into 50 µl of TOP10 electrocompetent cells from Invitrogen. 250 µl of SOC was added after electroporation, followed by a 1 hr incubation with shaking at 37°C. Thereafter, 10-50 µl of the transformation mix was plated on LA plates with 5ppm chloramphenicol (CMP) or LA plates with 5ppm CMP and 0.1ppm of cefotaxime (CTX) for selection of active BLA clones. The active BLA clones from the CMP + CTX plates were used for screening, whereas the random library clones from the CMP plates were sequenced to assess the quality of the library.

16 randomly chosen clones were sequenced. The clones contained different combinations of 1 to 7 mutations.

#### 20 D. Screen for improved expression

We found that when TOP10/pME27.1 is cultured in LB medium at 37C then the concentration of intact fusion protein peaks after one day and most of the fusion protein is degraded by host proteases after 3 days of culture. Degradation seems to occur mainly in the scFv portion of the CAB1 fusion protein as the cultures contain significant amounts of free BLA after 3 days, which can be detected by Western blotting, or nitrocefin (Oxoid, New York) activity assay. Thus we applied a screen to library NA05 that was able to detect variants of CAB1-scFv that would resist degradation by host proteases over 3 days of culture at 37 C.

Library NA05 was plated onto agar plates with LA medium containing 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime (Sigma). 910 colonies were transferred into a total of 10 96-well plates containing 100 ul/well of LA medium containing 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime. Four wells in each plate were inoculated with TOP10/pME27.1 as control and one well per plate was left as a blank. The plates were grown overnight at 37 C. The next day the cultures were used to inoculate fresh plates (production plates) containing

100 ul of the same medium using a transfer stamping tool and glycerol was added to the master plates which were stored at -70 C. The production plates were incubated in a humidified shaker at 37C for 3 days. 100 ul of BPER (Pierce, Rockford, IL) per well was added to the production plate to release protein from the cells. The production plate was  
5 diluted 100-fold in PBST (PBS containing 0.125% Tween-20) and BLA activity was measured by transferring 20 ul diluted lysate into 180 ul of nitrocephin assay buffer (0.1 mg/ml nitrocephin in 50 mM PBS buffer containing 0.125% octylglucopyranoside (Sigma)) and the BLA activity was determined at 490 nm using a Spectramax plus plate reader (Molecular Devices, Sunnyvale, CA).

10 Binding to CEA (carcinoembryonic antigen, Biodesign Intl., Saco, Maine) was measured using the following procedure: 96-well plates were coated with 100 ul per well of 5 ug/ml of CEA in 50 mM carbonate buffer pH 9.6 overnight. The plates were washed with PBST and blocked for 1-2 hours with 300 ul of casein (Pierce, Rockford, IL). 100 ul of sample from the production plate diluted 100-1000 fold was added to the CEA coated plate  
15 and the plates were incubated for 2 h at room temperature. Subsequently, the plates were washed four times with PBST and 200 ul nitrocephin assay buffer was added, and the BLA activity was measured as described above.

The BLA activity that was determined by the CEA-binding assay and the total BLA activity found in the lysate plates were compared and variants were identified which showed  
20 high levels of total BLA activity and high levels of CEA-binding activities.

The winners were confirmed in 4 replicates using a similar protocol: the winners were cultured in 2 ml of LB containing 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime for 3 days. Protein was released from the cells using BPER reagent. The binding assay was performed as described above but different dilutions of culture lysate were tested for each variant. Thus  
25 one can generate a binding curve which provides a measure of the binding affinity of the variant for the target CEA. Fig. 11A shows binding curves. Culture supernatants were also analyzed by SDS polyacrylamid electrophoresis. Fig. 11B shows the electropherogram of 7 variants from NA05. The band of the fusion protein is labeled for variant NA05.6. Table 6 shows a ranking of 6 variants. The data clearly show that NA05.6 produces significantly  
30 larger quantities of fusion protein compared to the fusion construct pME27.1.

Table 6 showing the sequence of 6 variants with the largest improvement in stability:

clone	mutations
NA05.6	R13K, T16G, W181V

NA05.8	R13K, F170Y, A234G
NA05.9	K3Q, S14P, L37V, E42G, E136Q, M146V, W181V, A234G
NA05.10	K3Q, L37V, P170Y, W181V
NA05.12	K3Q, S14P, L37V, M146V
NA05.15	M146V, F170Y, A194D

#### E. Construction of library NA06

Clone NA05.6 was chosen as the best variant and was used as the template for a second round of combinatorial mutagenesis. We used a subset of the same mutagenic primers that had been used to generate library NA05 to generate combinatorial variants with the following mutations: K3Q, L37V, E42G, E136Q, M146V, F170Y, A194D, A234G which had been identified in other winners from library NA05. We did not use the primer encoding mutation S14P as its sequence overlapped with mutations R13K and T16G that are present in NA05.6. A combinatorial library was constructed using QuikChange Multisite as described above and was called NA06. Template was pNA05.6 and 1  $\mu$ l of primers mix (10  $\mu$ M stock of all primers combined containing 1.25  $\mu$ M each primer) were used.

#### F. Screening of library NA06

The screen was performed as described above with the following modifications:

291 variants were screened on 3 96-well plates. 10  $\mu$ l sample from the lysate plates was added to 180  $\mu$ l of 10  $\mu$ g/ml thermolysin (Sigma) in 50 mM imidazole buffer pH 7.0 containing 0.005% Tween-20 and 10 mM calcium chloride. This mixture was incubated for 1 h at 37C to hydrolyze unstable variants of NA05.6. This protease-treated sample was used to perform the CEA-binding assay as described above.

Promising variants were cultured in 2 ml medium as described above and binding curves were obtained for samples after thermolysin treatments. Figure 11C shows binding curves for selected clones. It can be seen that a number of variants retain much more binding activity after thermolysin incubation than the parent NA05.6.

Table 7 showing 6 variants which are significantly more protease resistant than NA05.6:

Clone	mutations
NA06.2	R13K, T16G, W181V, L37V, E42G, A194D

NA06.4	R13K, T16G, W181V, L37V, M146V
NA06.6	R13K, T16G, W181V, L37V, M146V, K3Q
NA06.10	R13K, T16G, W181V, L37V, M146V, A194D
NA06.11	R13K, T16G, W181V, L37V, K3Q, A194D
NA06.12	R13K, T16G, W181V, L37V, E136Q

All 6 variants have the mutation L37V which was rare in randomly chosen clones from the same library. Further testing showed that variant NA06.6 had the highest level of total BLA activity and the highest protease resistance of all variants.

#### 5                    Example 4: Generation of an scFV that has pH-dependent binding

##### A.        Choosing positions for mutagenesis

The 3D structure of the scFv portion of NA06.6 was modeled based on the published crystal structure of a close homologue, MFE-23 [Boehm, M. K., A. L. Corper, T. Wan, M. K. Sohi, B. J. Sutton, J. D. Thornton, P. A. Keep, K. A. Chester, R. H. Begent and S. J. Perkins (2000) *Biochem J* 346 Pt 2, 519-28, Crystal structure of the anti-(carcinoembryonic antigen) single-chain Fv antibody MFE-23 and a model for antigen binding based on intermolecular contacts] using the software package MOE (Chemical Computing Group, Montreal, Canada) and using default parameters. A space filling model of the structure was visually inspected.

15 Side chains in the CDRs were ranked as follows: 0 = buried; 1 = partially exposed; 2 = completely exposed. Side chain distance to CDR3 was ranked as: 0 = side chain is in CDR3; 1 = side chain is one amino acid away from CDR3; 2 = side chain is two amino acids away from CDR3. In a few cases, residues flanking the CDRs were included if they fit the distance and exposure criteria.

20                Based on this ranking, the following side chains were targeted for mutagenesis:

a) exposure = 2 and distance = 2 or smaller

b) exposure = 1 and distance < 2

40 positions in the CDRs matched these criteria.

Fig. 14 shows the CDRs and the residues that were chosen for mutagenesis.

25                The table below shows the criteria and position of the 40 sites that were chosen for mutagenesis.

##### B.        Construction of library NA08

A combinatorial library was constructed where the 40 selected positions were randomly replaced with aspartate or histidine. The substitutions were chosen as it has been reported that ionic interactions between histidine side chains and carboxyl groups form the structural basis for the pH-dependence of the interaction between IgG molecules and the Fc receptor [Vaughn, D. E. and P. J. Bjorkman (1998) *Structure* 6, 63-73., Structural basis of pH-dependent antibody binding by the neonatal Fc receptor].

The QuikChange multi site-directed mutagenesis kit (QCMS; Stratagene Catalog # 200514) was used to construct the combinatorial library NA08 using 40 mutagenic primers. The primers were designed so that they had 17 bases flanking each side of the codon of interest based on the template plasmid NA06.6. The codon of interest was changed to the degenerate codon SAT to encode for aspartate and histidine. All primers were designed to anneal to the same strand of the template DNA (i.e., all were forward primers in this case). The QCMS reaction was carried out as described in the QCMS manual with the exception of the primer concentration used; the manual recommends using 50-100ng of each primer in the reaction, whereas significantly lower amounts of each primer were used in this library as this results in a lower parent template background. In particular, 0.4 $\mu$ M of all primers together were used. The individual degenerate primer concentration in the final reaction was 0.01  $\mu$ M (approximately 2.5ng).

The QCMS reaction contained 50-100 ng template plasmid (NA06.6, 5178bp), 1  $\mu$ l of primers mix (10 $\mu$ M stock of all primers to give the desired primer concentration mentioned above), 1  $\mu$ l dNTPs (QCMS kit), 2.5  $\mu$ l 10x QCMS reaction buffer, 18.5  $\mu$ l deionized water, and 1  $\mu$ l enzyme blend (QCMS kit), for a total volume of 25  $\mu$ l. The thermocycling program was 1 cycle at 95° for 1 min., followed by 30 cycles of 95°C for 1 min., 55°C for 1 min., and 65°C for 10 minutes. *DpnI* digestion was performed by adding 1  $\mu$ l *DpnI* (provided in the QCMS kit), incubation at 37°C for 2 hours, addition of 0.5  $\mu$ l *DpnI*, and incubation at 37°C for an additional 2 hours. 1  $\mu$ l of each reaction was transformed into 50  $\mu$ l of TOP10 electrocompetent cells from Invitrogen. 250  $\mu$ l of SOC was added after electroporation, followed by a 1 hr incubation with shaking at 37°C. Thereafter, 10-50  $\mu$ l of the transformation mix was plated on LA plates with 5ppm chloramphenicol (CMP) or LA plates with 5ppm CMP and 0.1ppm of cefotaxime (CTX) for selection of active BLA clones. The number of colonies obtained on both types of plates was comparable (652 on the CMP plate and 596 colonies on the CMP + CTX plate for 10  $\mu$ l of the transformation mix plated). Active BLA clones from the CMP + CTX plates were used for screening, whereas random library clones from the CMP plates were sequenced to assess the quality of the library.



Primers for the reaction are shown in Table 8.

Table 8 Primers for CDRs:

residue	CDRs	position	exposure	distance to CDR3	primer sequence
K		30	2	2	cttctggcttcaacattsatgactcctatatgcactg
D	H1	31	2	1	ctggcttcaacattaaasattcctatatgcactgggt
S	H1	32	1	1	gcttcaacattaaagacsattatatgcactgggtgag
Y	H1	33	2	1	tcaacattaaagactccsatatgcactgggtgaggca
H	H1	35	1	1	ttaaagactcctatatgsattgggtgaggcaggggcc
W	H2	50	2	1	gcctggagtggattggasatattgatcctgagaatgg
D	H2	52	2	2	agtggattggatggattsatcctgagaatggtgatac
E	H2	54	2	2	ttggatggattgatcctsataatggtgatactgaata
N	H2	55	2	2	gatggattgatcctgagsatggtgatactgaatatgc
D	H2	57	2	1	ttgatcctgagaatggtsatactgaatatgccccgaa
T	H2	58	1	1	atcctgagaatggtgatsatgaatatgccccgaagtt
E	H2	59	2	1	ctgagaatggtgataactsattatgccccgaagtcca
P	H2	62	2	1	gtgatactgaatatgccsataagttccagggcaaggc
K	H2	63	2	3	atactgaatatgccccgsatttccagggcaaggccac
Q	H2	65	2	2	aatatgccccgaagttcsatggcaaggccacttttac
E		98	1	0	ccgtctattattgtaatsatgggactccgactggggcc
G		99	1	0	tctattattgtaatgagsatactccgactgggcccgt
T	H3	100	2	0	attattgtaatgaggggsatccgactgggcccgtacta
P	H3	101	2	0	attgtaatgaggggactsatactgggcccgtactactt
T	H3	102	2	0	gtaatgaggggactccgsatgggcccgtactactttga
G	H3	103	2	0	atgaggggactccgactsatccgtactactttgacta
P	H3	104	2	0	aggggactccgactgggsattactactttgactactg
Y	H3	106	2	0	ctccgactgggcccgtacsattttgactactggggcca
S	L1	162	2	2	taacctgcagtgccagcsatagtgttaagttacatgca
S	L1	163	2	1	cctgcagtgccagctcasatgttaagttacatgcactg
V	L1	164	1	1	gcagtgccagctcaagtsatagttacatgcactgggt
S	L1	165	2	1	gtgccagctcaagtgtasattacatgcactgggtcca
Y	L1	166	2	1	ccagctcaagtgttaagtsatagcactgggtccagca
Y		183	1	0	ctcccaaactcgtgattsatagcacatccaacctggc

S	L2	184	2	0	ccaaactcgtgatttatsatacatccaacctggcttc
T	L2	185	1	1	aactcgtgatttatagcsattccaacctggcttctgg
S	L2	186	2	2	tcgtgatttatagcacasataacctggcttctggagt
N	L2	187	2	1	tgatttatagcacatccsatctggcttctggagtccc
A	L2	189	1	1	atagcacatccaacctgsattctggagtcctgctcg
S	L2	190	2	1	gcacatccaacctggctsatggagtcctgctcgctt
R	L3	225	2	2	cttattactgccagcaasattctagttaccactcac
S	L3	226	2	2	attactgccagcaaagasatagttaccactcacgt
S	L3	227	1	2	actgccagcaaagatctsattaccactcacgttcg
Y	L3	228	1	2	gccagcaaagatctagtsatccactcacgttcggtg
L	L3	230	1	2	aaagatctagttaccasatacgttcggtgctggcac

### C. Sequencing of variants

5 Variants were grown overnight with shaking at 37°C in 5mL cultures of LA containing 5ppm of CMP. Miniprep DNA was prepared using a Qiagen kit and the BLA gene within each clone was sequenced using the M13 reverse and nsal54f primers.

M13 reverse: CAGGAAACAGCTATGAC

nsal54f: GGACCACGGTCACCGTCTCCTC

### D. Screen pH-dependent binding

10 Library NA08 was plated onto agar plates with LA medium containing 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime (Sigma). 552 colonies were transferred into a total of six 96-well plates containing 100 ul/well of LA medium containing 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime. Four wells in each plate were inoculated with TOP10/NA06.6 as a reference. The plates were grown overnight at 37 C. The next day the  
15 cultures were used to inoculate fresh plates (production plates) containing 100 ul of the same medium using a transfer stamping tool and glycerol was added to the master plates which were stored at -70 C. The production plates were incubated in a humidified shaker at 37C for 2 days. 100 ul of BPER (Pierce, Rockford, IL) per well was added to the production plates to  
20 release protein from the cells. The production plates were diluted 100-fold in PBST (PBS containing 0.125% Tween-20) and BLA activity was measured by transferring 20 ul diluted lysate into 180 ul of nitrocefin assay buffer (0.1 mg/ml nitrocefin in 50 mM PBS buffer

containing 0.125% octylglucopyranoside (Sigma)) and the BLA activity was determined at 490 nm using a Spectramax plus plate reader (Molecular Devices, Sunnyvale, CA).

Binding to CEA (carcinoembryonic antigen, Biodesign Intl., Saco, Maine) was measured using the following procedure: 96-well plates were coated with 100 ul per well of 5 ug/ml of CEA in 50 mM carbonate buffer pH 9.6 overnight. The plates were washed with PBST and blocked for 1-2 hours with 300 ul of casein (Pierce, Rockford, IL). 100 ul of sample from the production plate diluted 100-1000 fold was added to the CEA coated plate and the plates were incubated for 2 h at room temperature. Subsequently, the plates were washed four times with PBST and 200 ul nitrocefin assay buffer was added, and the BLA activity was measured as described above. CEA binding was measured in 50 mM phosphate buffer pH 6.5 and in a separate experiment in 50 mM phosphate buffer pH 7.4.

The BLA activity that was determined by the CEA-binding assay at pHs of 6.5 and 7.4, and the total BLA activity found in the lysate plates were compared and variants were identified which showed good binding to CEA at pH 6.5 but significantly weaker binding at pH 6.5. A comparison of the binding at pH6.5 versus pH 7.4 is shown in Fig. 13.

Winners were confirmed by culturing them in 5 ml of LB medium containing 5 mg/l chloramphenicol and 0.1 mg/l cefotaxime (Sigma) for 2 days at 37 C. Subsequently, the cultures were centrifuged and the pellet was suspended in 375 ul of BPER reagent to release the fusion protein. The BLA activity in each sample was determined by transferring 20 ul of the appropriately diluted sample to 180 ul of 180 ul of nitrocefin assay buffer (0.1 mg/ml nitrocefin in 50 mM PBS buffer containing 0.125% octylglucopyranoside (Sigma)) and the absorbance at 490 nm was monitored. One unit of activity was defined as the amount of BLA that leads to an absorbance increase of one mOD per minute. The samples were diluted based on their total content of BLA activity and the CEA-binding assay was performed as described above but adding various sample dilutions to each well.

Thus, one can obtain binding curves for each sample that reflect the affinity of the variants to CEA. Figure 15 shows CEA-binding curves measured at pH 7.4 and pH 6.5 for several variants of interest. All 5 variants show increased pH-dependence of CEA binding. Whereas, the parent NA06.6 binds only slightly better at pH 6.5 compared to pH 7.4, some of the variant show much stronger binding to CEA at pH 6.5 compared to pH 7.4. Of particular interest are variants NA08.15 and NA08.17 which show very weak binding to CEA at pH 7.4 but significant binding at pH 6.5.

Table 9, below, shows variants with the greatest binding improvement.

Table 9:

clone	mutations
NA08.1	W50H, Y166D
NA08.3	S190D, S226D
NA08.4	S190D, T100D
NA08.9	Y166D
NA08.12	T102H, Y166D, S226D
NA08.13	Q65H, S184D, S226D
NA08.14	P101D
NA08.15	S184D, S226D
NA08.17	S184D, W50H
NA08.24	T102D, S226D
NA08.45	T102D, Y166D
NA08.51	P104H, Y166D
NA08.64	Q65D, Y166D

#### Example 5: purification of ME27.1

Purification of ME27.1 from cell extract was done using Cation Exchange Chromatography. This was performed with the aid of a high performance liquid chromatographic system (AKTA™explorer 10, Amersham Biosciences) on a 7.3mL CM Ceramic HyperF cation exchange column (Ciphergen Biosystems). The column was first equilibrated with loading/equilibration buffer. The prepared ME27.1 extract was applied to the column at 300cm/hr, followed by washing with the equilibration buffer. The bound proteins were eluted using a sodium chloride gradient. The eluted fractions were analyzed using colorimetric activity assay (o-nitrocefan as substrate) and 4-12% Bis-Tris SDS-PAGE reducing gel with MES running buffer (Novex).

#### Buffers

The following buffers were used for the purification using Cation Exchange Chromatography:

Loading/Equilibration: 50mM Sodium Acetate, pH 5

Elution: 50mM Sodium Acetate containing 1M Sodium Chloride, pH 5

Regeneration: 0.5M Sodium Hydroxide and 1M Sodium Acetate, pH 4

#### a) Extract Sample Preparation

*E. coli* cells containing gene of interest were cultured in TB broth. The production media flasks were incubated at 30° C, 150-200 rpm for 18-24 hours. After fermentation, the *E. coli* cells were centrifuged at 4,000 rpm for 30 minutes. The supernatant was discarded and BPER detergent (Pierce product #78266) was added to the pellet to lyse the cells (27 mL of B-PER per gram of cell pellet). The lysed cells were centrifuged at 18,000 rpm for 20 minutes to remove cell debris. The supernatant containing the *ME27.1*, referred to as the “extract”, was used for the subsequent purification experiments. The extract was stored at 4°C until use for subsequent purification experiments.

- 10 The following sample pretreatment steps were followed to prepare the extract for subsequent chromatography purification experiments.

Dilute *ME27.1* extract (conductivity 9.5mS/cm and pH 7.3) with 1 part of equilibration buffer (see below). pH of the diluted extract was 6.51.

- 15 Adjust pH to 5.0 using ~20% acetic acid.

Filter pH adjusted extract through 0.2µm filter unit with ~1% diatomaceous earth as admix.

- 20 Figure 16 shows the overall the chromatogram obtained using the 18 hours old extract. The fractions from peak #1 constitute 23% of the total eluted activity, with the major protein band near molecular weight (MW) equivalent to the *ME27.1* MW (see Figure 17, gel on left). Peaks #2, #3 and #4 constitute 76.6% of the total eluted activity but SDS-PAGE gel (see Figure 17, gel on right) shows that these fractions contains relatively small amount of protein near *ME27.1* MW. There was an increasing amount protein bands at MW below *ME27.1* MW.

- 25 Figure 18 shows the overall chromatogram obtained using 26 hours old extract. The chromatogram is significantly different from Figure 16. The relative proportion of peak #1 to peak #2 has decreased significantly for the 26 hours extract compared to the 18 hour extract. SDS-PAGE gels shows that the small peak #1 contains mostly protein near *ME27.1* MW but the remainder fractions (from peak #2 and onwards) contain mostly protein bands with MW lower than *ME27.1* (see Figure 19).

Figure 20 shows the overall chromatogram obtained using >> 26 hours old (4-5days) extract. The chromatogram is significantly different from Figure 16 and Figure 18. The distinct peak #1 found in 18 hrs (Figure 16) and 26 hours (Figure 18) old extract has collapsed into a

shoulder of equivalent peak #2 found in Figure 16 and Figure 18. SDS-PAGE gels shows that the shoulder contains two main bands, one near the ME27.1 MW and the other at lower MW. The main peak #1 fractions contain 88% of the activity eluted, and yet the main protein band is below the MW of ME27.1. Mass spec analysis confirmed that the lower band is degraded ME27.1 (see Figure 21, left gel circled band).

Purification of NA05.6 extract was done using Anion Exchange Chromatography follow by Affinity chromatography using aminophenylboronic acid (PBA) resin.

The anion exchange chromatography was performed with the aid of a high performance liquid chromatographic system (AKTA™explorer 10, Amersham Biosciences) on a 7mL Poros HQ anion exchange column (PE Biosystems). The column was first equilibrated with loading/equilibration buffer. The prepared NA05.6 extract was applied to the column at 300cm/hr. The NA05.6 was collected as the flow through and wash fractions. These fractions were analyzed for activity using colorimetric assay (o-nitrocefan as substrate) and purity using 4-12% Bis-Tris SDS-PAGE reducing gel with MES running buffer (Novex).

The PBA step was done using a 5 mL column containing PBA resin (Sigma A-8530 m-aminophenylboronic acid resin). The column was first equilibrated with 20mL of each equilibration buffers by gravity flow. The anion exchange partially purified flow through was used as the PBA feed. The feed was applied to the column by gravity flow and wash with wash buffer. The bound sample was eluted with elution buffer. Samples were analyzed for activity using colorimetric assay (o-nitrocefan as substrate) and purity using 4-12% Bis-Tris SDS-PAGE reducing gel with MES running buffer (Novex).

### Buffers

The following buffers were used for the purification using Anion Exchange Chromatography:

Load/ Equilibration: 50mM Tris, pH 7.4

Regeneration: 0.5M NaOH

The following buffers were used for the purification using PBA Affinity Chromatography:

Equilibration Buffers: 0.5 M sorbitol with 1 M NaCl; 0.5 M Borate at pH 7; and 20 mM

TEA with 0.5 M NaCl, pH 7

Wash buffer: 20mM TEA with 0.5M NaCl

Elution Buffer: 0.5 M borate with 0.5M NaCl

b) Extract Sample Preparation for Anion Exchange  
Chromatography

5

The flow through contains 79% of the activity loaded onto the anion exchange column. Subsequent purification of the flow through fraction using PBA affinity chromatography shows that all the protein eluted from contain protein near NA05.6 MW (*See* Figure 22) and 80% of the loaded activity was recovered. This indicates that all the NA05.6 molecule is intact.

10

Various publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

**WHAT IS CLAIMED IS:**

1. A method of improving a binding characteristic of a binding sequence for a target  
5 comprising:

a) contacting the target with a reporter fusion under conditions that allow the reporter fusion to bind to the target, wherein the reporter fusion comprises a reporter sequence and a variant binding sequence that is derived from a prototype binding sequence that binds the target, and

10 b) selecting the reporter fusion if it has an improved binding characteristic compared to the prototype binding sequence.

2. The method of Claim 1 wherein the selected reporter fusion binds to the target with an affinity that is greater than the binding affinity of the prototype binding sequence for the  
15 target.

3. The method of Claim 1 wherein the selected reporter fusion binds to the target with a greater specificity than the prototype binding sequence has for the target.

20 4. The method of Claim 1, wherein step (b) comprises incubating the reporter fusion in the presence of proteases and/or under conditions which degrade or destabilize the reporter fusion.

25 5. The method of Claim 4, wherein the conditions are at least one of heat, pH or incubation in the presence of solutes that affect stability.

6. The method of Claim 1 further comprising repeating steps (a) and (b) one or more times, wherein the binding sequence of the reporter fusion selected in a previous step (b) is the prototype binding sequence of the subsequent step (a).

30 7. The method according to Claim 1 further comprising removing the reporter sequence from the binding sequence of the reporter fusion selected in step (b).



8. A method of improving a binding characteristic of a binding sequence for a target comprising:

- a) contacting a target with a library comprising a multiplicity of reporter fusions, under conditions that allow a reporter fusion to bind the target, wherein said reporter fusions comprise a reporter sequence and a variant binding sequence derived from a prototype binding sequence that binds the target, and
- b) selecting a reporter fusion bound to the target that has an improved binding characteristic.

9. The method of Claim 8 wherein the selected reporter fusion binds to the target with an affinity that is greater than the binding affinity of the prototype binding sequence for the target.

10. The method of Claim 8 wherein the selected reporter fusion binds to the target with a greater specificity than the prototype binding sequence has for the target.

11. The method of Claim 8 further comprising repeating steps (a) and (b) one or more times, wherein the binding sequence of the reporter fusion selected in a previous step (b) is the prototype binding sequence of the subsequent step (a).

12. The method according to Claim 8 further comprising removing the reporter sequence from the binding sequence of the reporter fusion selected in step (b).

13. A method of improving the binding affinity of a binding sequence for a target comprising:

- a) making a reporter fusion by covalently linking a prototype binding sequence to a reporter sequence,
- b) modifying the binding sequence to produce a variant binding sequence,
- c) contacting the target with the reporter fusion under conditions that allow the reporter fusion to bind the target, and
- d) selecting the reporter fusion if it has an improved binding characteristic.

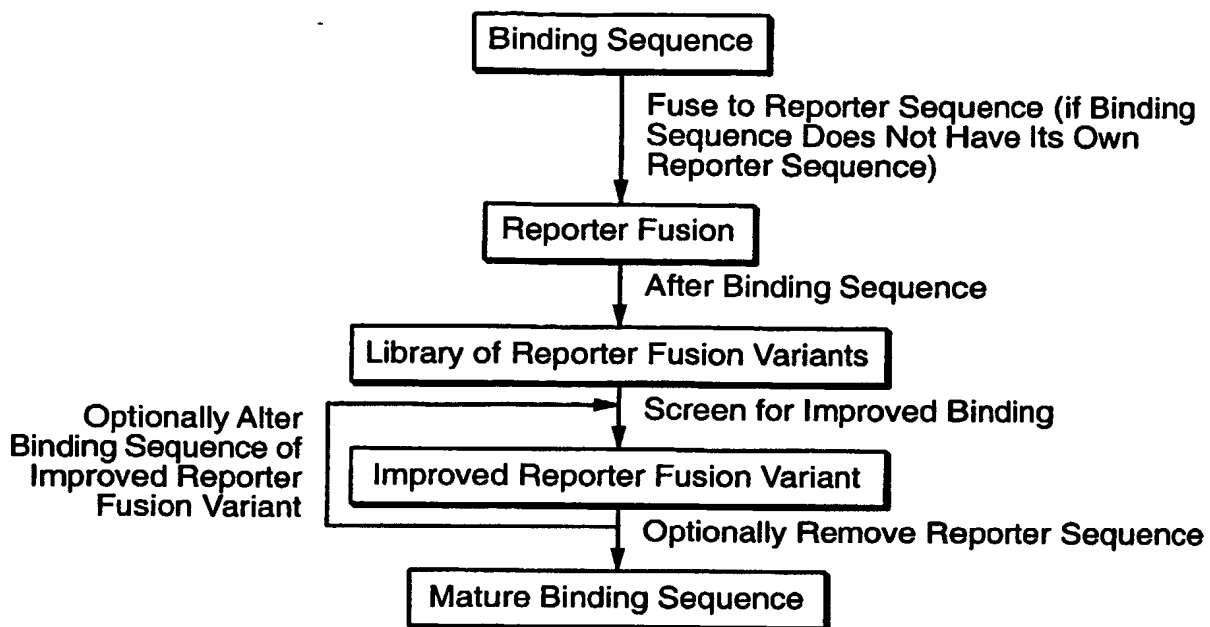
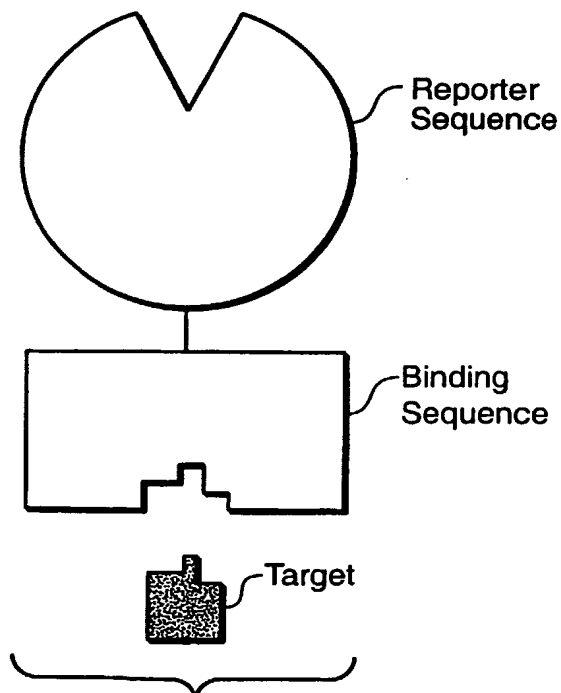
14. The method of Claim 13 wherein the selected reporter fusion binds to the target with an affinity that is greater than the binding affinity of the prototype binding sequence for the target.

5           15. The method of Claim 13 wherein the selected reporter fusion binds to the target with a greater specificity than the prototype binding sequence has for the target.

10           16. The method according to Claim 13 further comprising repeating steps (a) through (d) one or more times, wherein the variant binding sequence of a reporter fusion selected in a previous step (d) is the prototype binding sequence of the subsequent step (a).

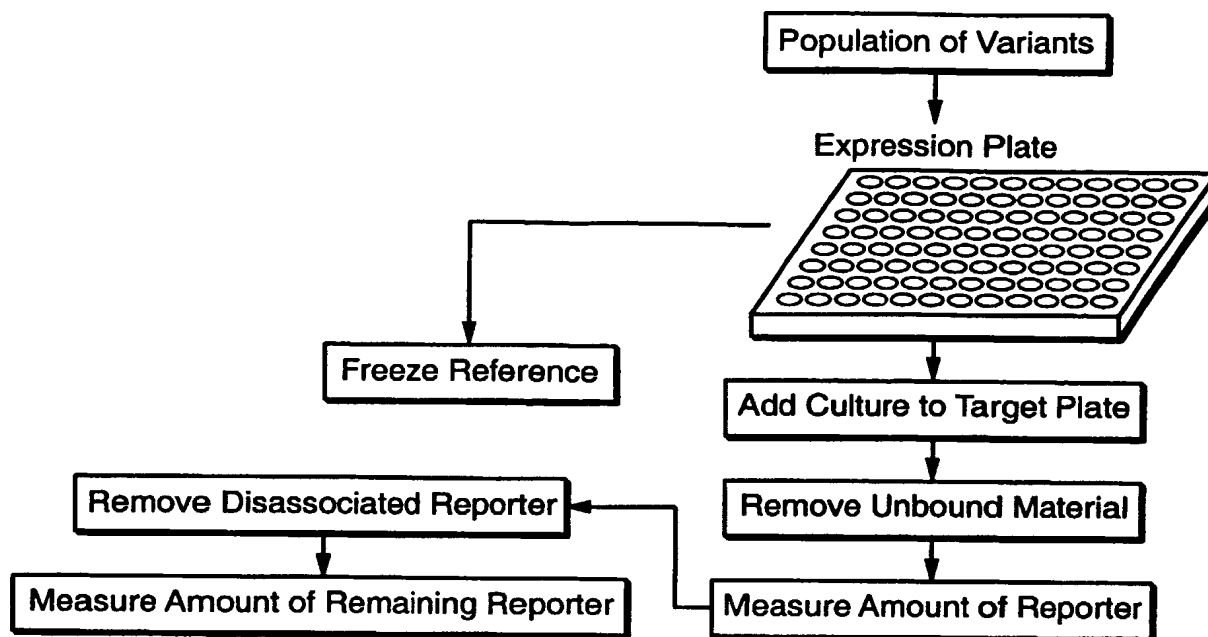
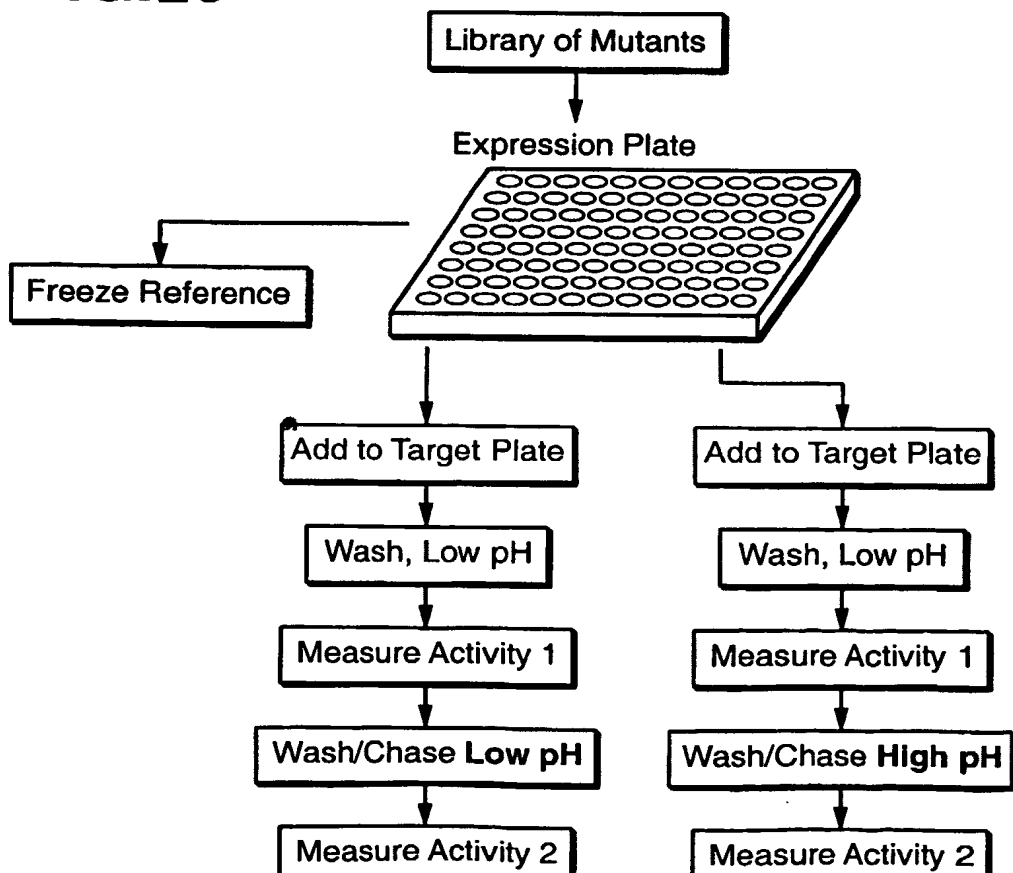
17. The method according to Claim 13 further comprising removing the reporter sequence from the binding sequence of the reporter fusion selected in step (d).

1 / 28

**FIG.\_1****Definitions****FIG.\_2**

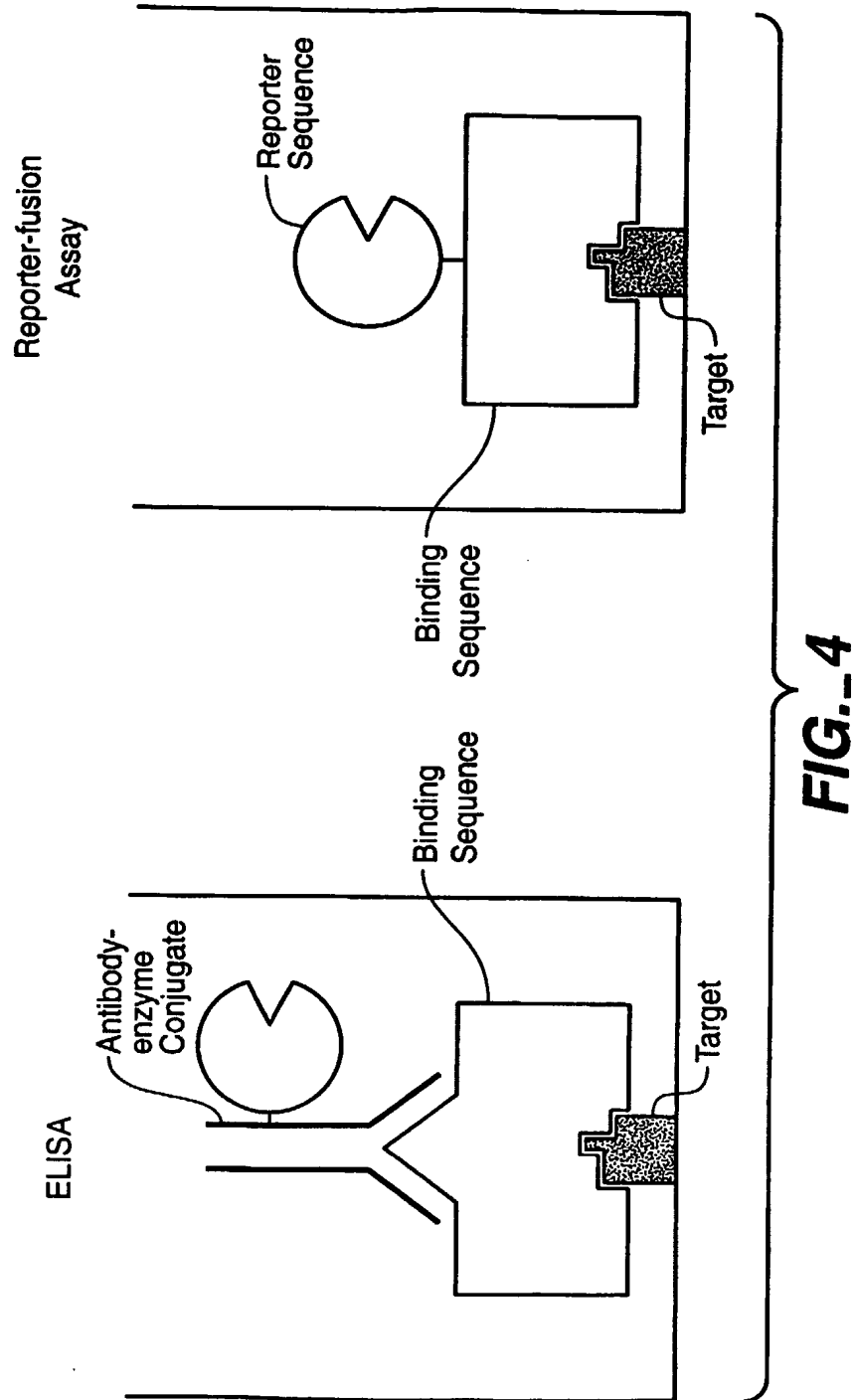
SUBSTITUTE SHEET (RULE 26)

2 / 28

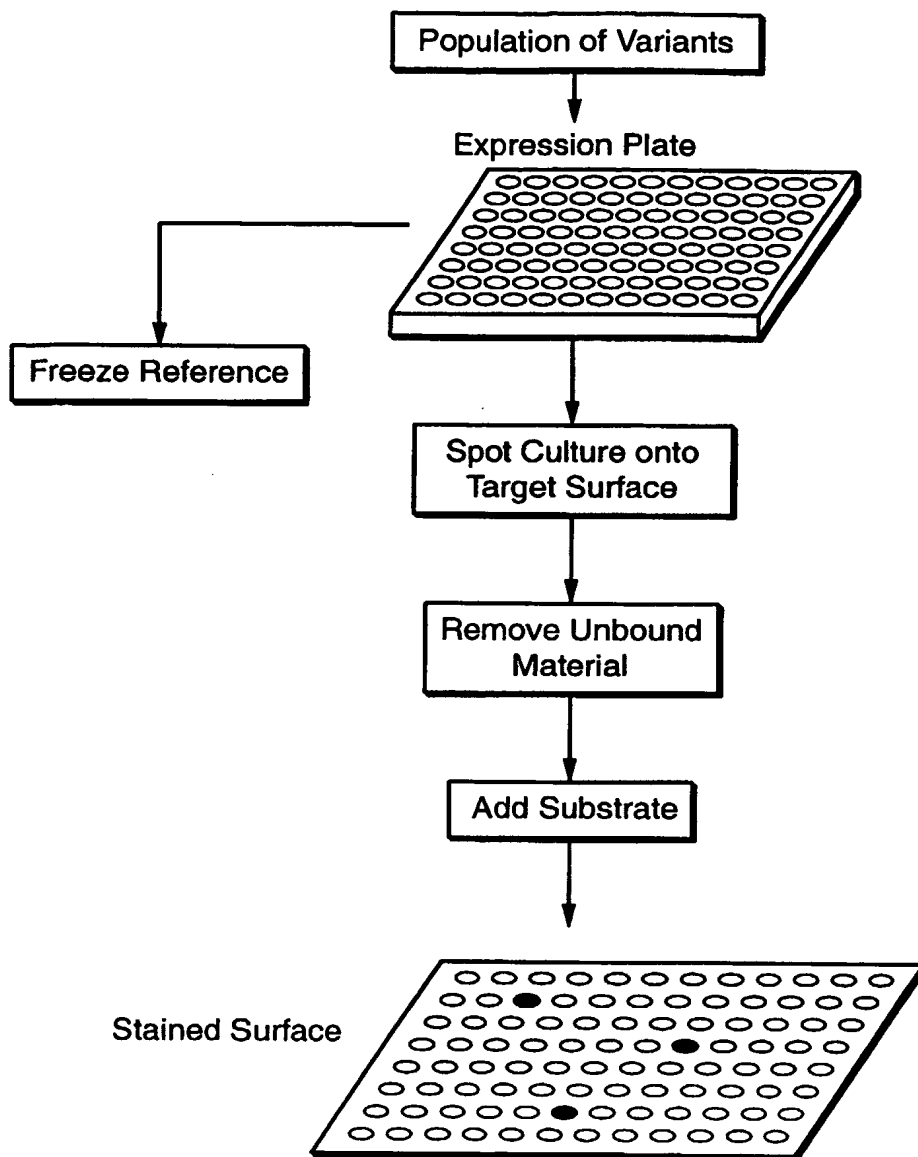
**FIG.\_3****Screening for Slow  $k_{off}$** **FIG.\_5****Screening for pH-dependent  $k_{off}$** 

3 / 28

Comparison Between a Traditional ELISA Assay and the Assay of a Reporter Fusion



4 / 28

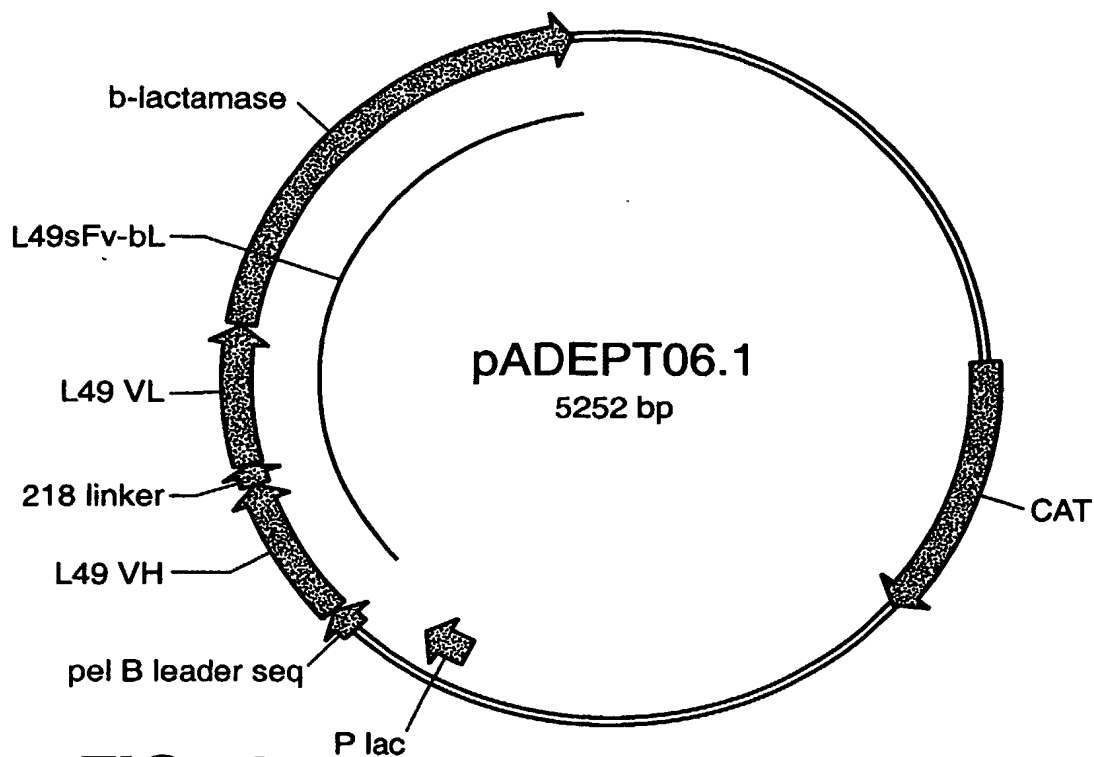
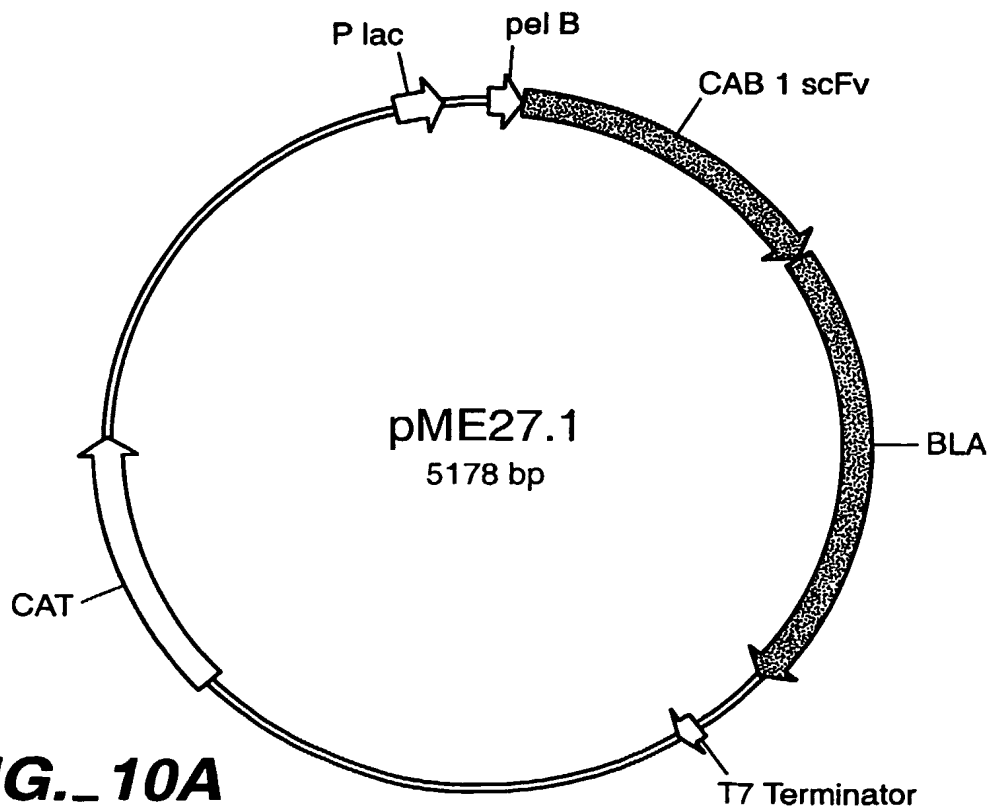
**Screening via Spotting Samples onto Target Surface****FIG.\_6**

5 / 28

10 20 30 40  
| | | |  
TPVSEKQLAE VVANTITPLM KAQSVPGMAV AVIYQGKPHY  
50 60 70 80 90 100  
| | \* \* | | | |  
YTFGKADIAA NKPVTPQTLF ELGSISKFT GVLGGDAIAR GEISLDDAVT RYWPQLTGKQ  
110 120 130 140 150 160  
| | | | \* \* | |  
WQGIRMLDLA TYTAGGLPLQ VPDEVTDNAS LLRFYQNWQP QWKPGTTRL Y ANASILGFGA  
170 180 190 200 210 220  
| | | | | |  
LAVKPSGMPY EQAMTTRVLK PLKLDHTWIN VPKAEAAHYA WGYRDGKAVR VSPGMLDAQA  
230 240 250 260 270 280  
| | | | | |  
YGVKTNVQDM ANWVMANMAP ENVADASLKQ GIALAQSR YW RIGSMYQGLG WEMLNHPVEA  
290 300 310 320 330 340  
| | | \*\* \* | | |  
NTVVEGSDSK VALAPLPVAE VNPPAPPVKA SWVHKTGSTG GFGSYVAFIP EKQIGIVMLA  
350 360  
| |  
NTSYPNPARV EAAYHILEAL Q

**FIG.\_7**

6 / 28

**FIG. 8****FIG. 10A**

SUBSTITUTE SHEET (RULE 26)



7 / 28

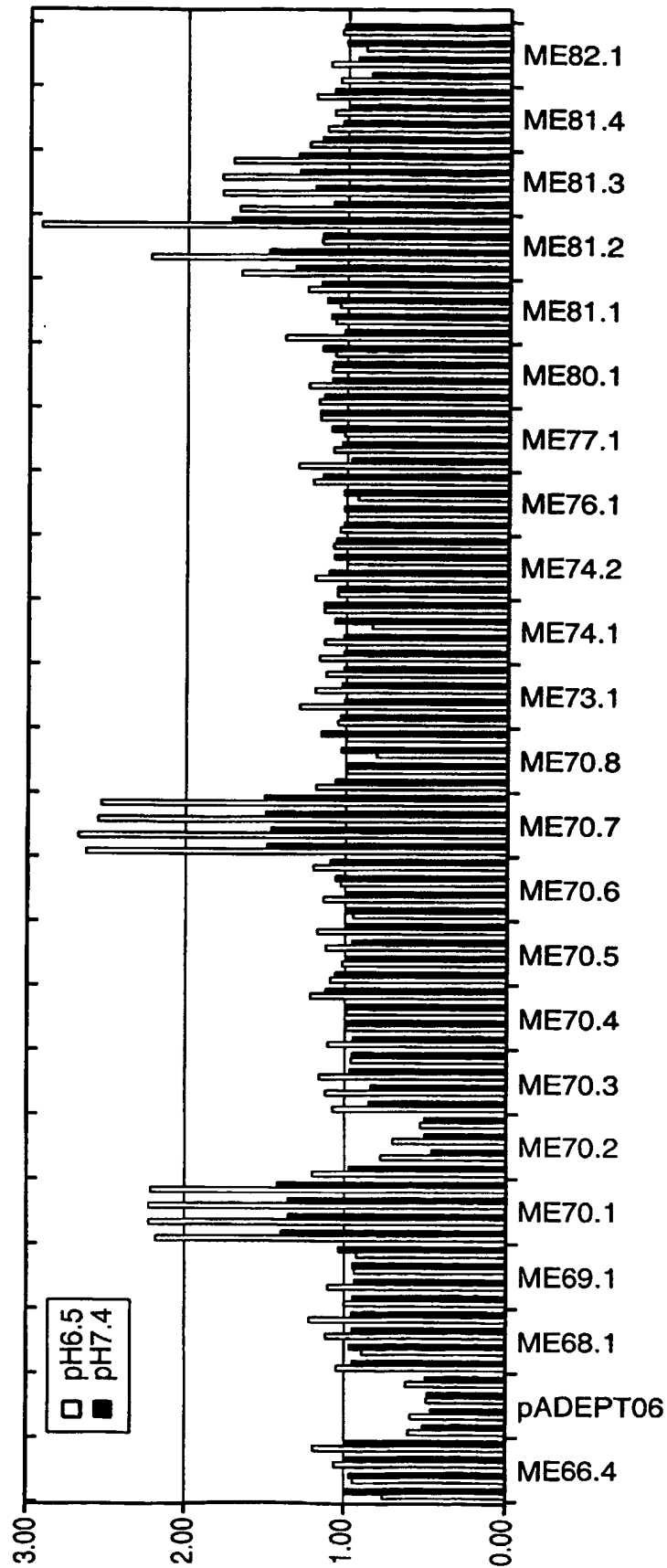


FIG.\_9



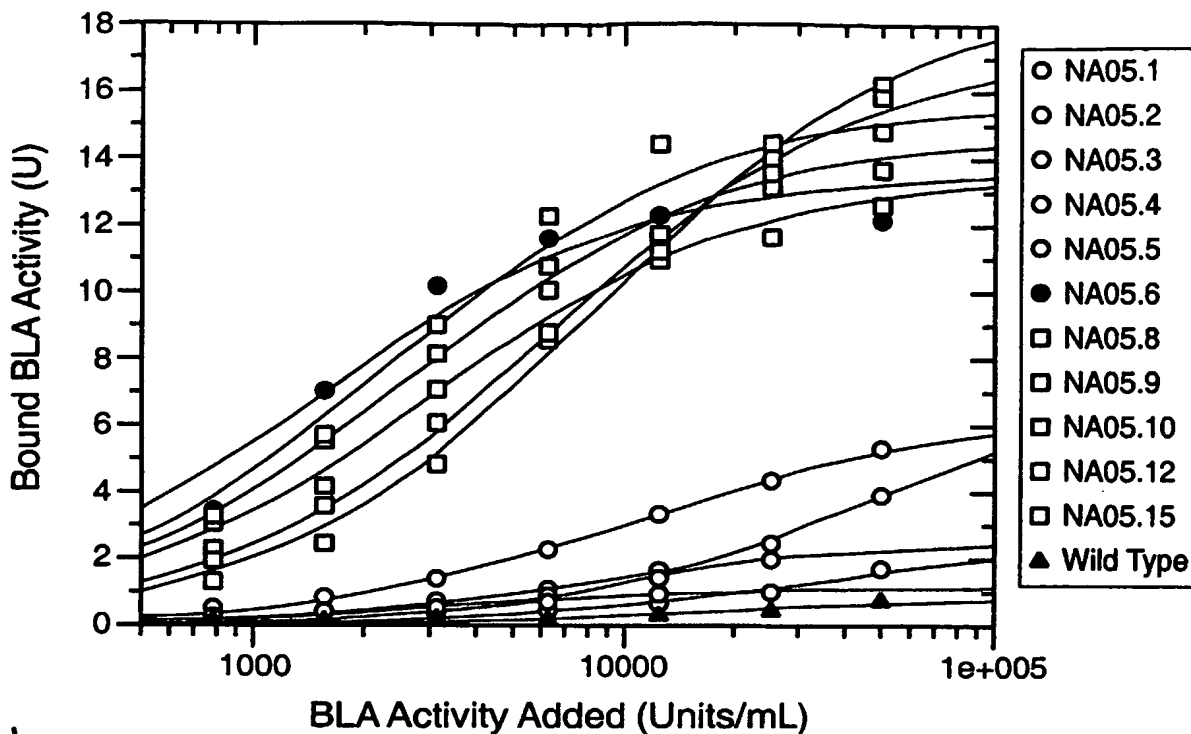
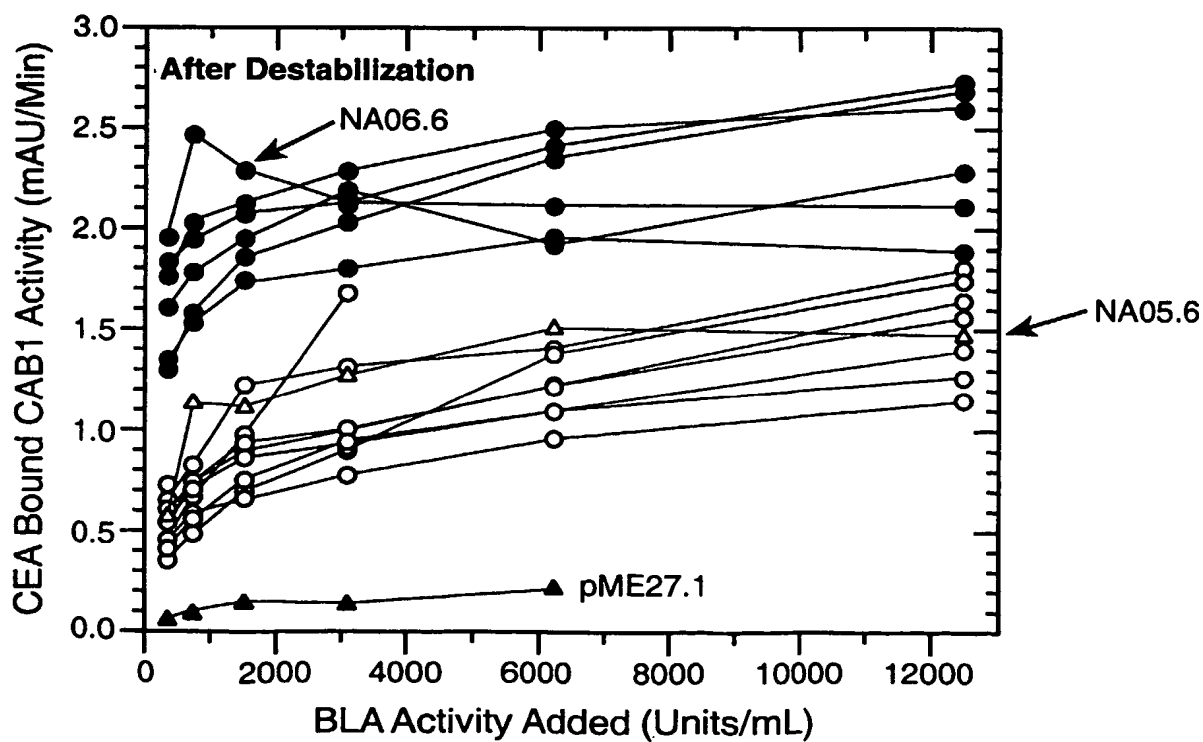


cataccgaaacgcgcgagacgaaagggccctcgtgatacgcctatcttataaggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgggaaacccctatttgtttattttctaaatacatctcaaatatgtatccgctcatgagacaataaacctgtggcagcatcacccgacgcactttgcgcggaataaaataacctgtgacgggaagatcacttcgcagaataaaataacctgtgtccctgttgataccgggaagccctgggccaacttttggcgaaatgagacgttgatcggcacgtaaagaggttccaactttaccataatgaaataagatcactaccggcgctattttttgagttatcgagattttcaggagctaagggaagctaaaaatggagaaaaaaatcactgggatataccaccggttgatatatcccaatggcatctgtaagaaacatttttgaggcatcttcagtcaggtgtcactataaccagaccggttcagctggataattacggcctttttaagaccgttaagaaaaataagcacagaattttatccggcctttattcacattcttgcgcgcctgatgaatgctcatcgggaattccgtatggcaatgaaagacggtgagctggtgatatgggatatgtttcacccctgtttacaccggttttccatgagcaaaactgaaacggttttcatcgtcttgagtgaaataccacgacgatttccggcagtttctacacatatattcgcgaagatgtggcgtgttacgggttaaaaactggcctatttccctaagggtttattgagaatatgttttcgtctcagccaatccctgggtgagtttcaccagtttttgatttaaacgtggccaatatgacaaacttcttcgcccccggttttcacgatgggcaaatattatcagcaagggcacaaggtgtgatgcgctggcgattcaggttcatcatgcgcgtctggtgatggcttccatgtcggcagaatgcttaatgaattacaacagctactcgcgatgagtggcagggcggtgtaagacagatcgcctgagataggtgccctgactgattaagcatgtgtaactgtcagaccaagtttactcatatatacttttagatttgatttaaaaacttcattttttaattaaaaaggatctagtggaagatcccttttgataatctcgtctgcaaacaaaaaacccagctaccagcgtgtgtttgtttgcggatcaagagctaccaactctttttccgaaaggttaactggcgagcaagacataccaaatactgttcttctagtgtagccgtagtttaggccaccacttcaagaaactctgtagcaccgcctacatacctcgcctctgcgaatcctgttaccagtggtcgtgccagtggtgataaagtcgttcttaccgggttggactcaagacgatatgttaccggataaggcgacggtcgggttgaaacgggggttcgtgcacacagcccagcttggagcgaacgcactacaccgaaactgagatacctacagcgtgagctatgagaaagcgccacgttcccccgaagggagaaaggcgacaggtatccggtaaaggcgaggtcggaaacgagagcgcacgagggaggttccaggggggaaaaagccttagtatctttatagtcctgtcgggtttcggcacctctgacttgagcgtcgatttttgtgatgctcgtcagggggggcgagcctatggaaaaacgcgcagcaacgggccttttacgggttccctggccttttgtcgccttttgtcacatgttcttccctggttatccctgattctgtggataacggtattaccgcctttgagtgagctgataccgctcgcgcagccgaacgacgagcgcagcgcagtcagtgagcgggaagcgggaagagcgcccaatacgcgaacacgcctctcccgcggttgccgattcattaatgcagctggcacgcagaggtttcccgactggaaagcggcgagtgagcgaacgcaattaatgtgagttagtcatacattagggcaccccggttttacactttatgcttcggctcgtatgttgtggaattgtgagcgggataaacaatttcacacaggaacacagctatgaccatgattacggccaagctatttaggtgacactatagaataactcaagcttttctagattaaagg

**FIG. 10C-2**

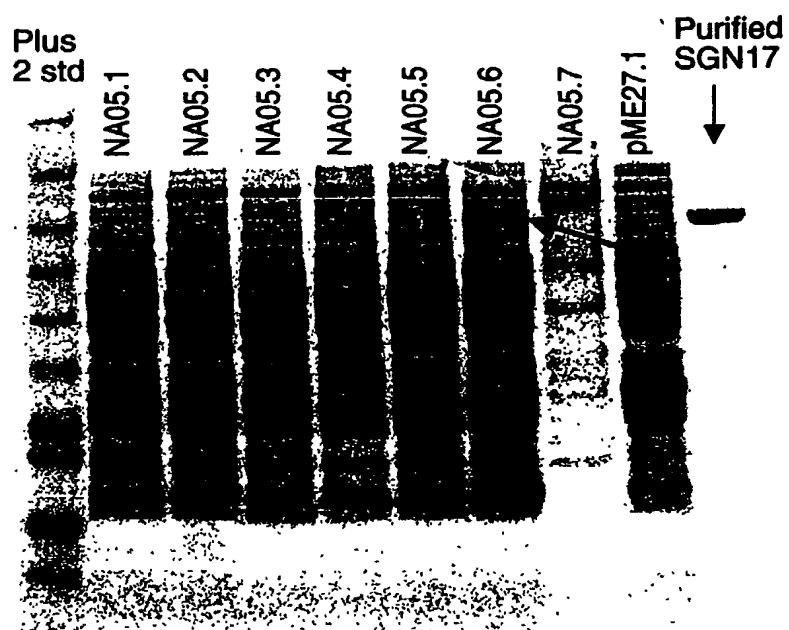
26)

11 / 28

**FIG. 11A****FIG. 11C**

SUBSTITUTE SHEET (RULE 26)

12 / 28

**FIG. 11B**

13 / 28

Pos. Heavy Chain	Number of Observations	Observed Frequencies of 5 Most Abundant Amino Acids in Alignment of Human Sequences												CAB1 Sequence	CDR	Mutated Residues
		E	Q	D	G	A	L	I	V	K	R	S	T			
1	291	E	0.616	Q	0.346	D	0.014	G	0.014	A	0.003	L	0.003	Q		
2	293	V	0.887	M	0.027	L	0.024	S	0.020	I	0.017	A	0.007	V		
3	291	Q	0.852	H	0.034	R	0.027	T	0.027	E	0.014	V	0.014	K		1
4	282	L	0.975	V	0.011	A	0.007	D	0.004	M	0.004			L		
5	276	V	0.645	Q	0.148	L	0.120	R	0.022	M	0.014	N	0.014	Q		
6	267	E	0.693	Q	0.263	A	0.022	D	0.011	G	0.007	R	0.004	Q		
7	265	S	0.951	W	0.019	X	0.015	T	0.008	A	0.004	N	0.004	S		
8	266	G	0.989	S	0.008	T	0.004							G		
9	274	G	0.624	A	0.193	P	0.164	S	0.011	E	0.004	H	0.004	A		
10	271	G	0.638	E	0.192	D	0.081	A	0.070	T	0.011	V	0.007	E		
11	270	L	0.681	V	0.270	F	0.030	S	0.019					L		
12	267	V	0.757	K	0.154	I	0.026	N	0.022	L	0.015	A	0.007	V		
13	247	K	0.474	Q	0.428	R	0.049	E	0.034	G	0.004	H	0.004	R		1
14	251	P	0.968	A	0.012	K	0.008	G	0.004	L	0.004	S	0.004	S		1
15	244	G	0.783	S	0.156	T	0.033	P	0.016	K	0.008	E	0.004	G		
16	243	G	0.488	E	0.131	Q	0.107	A	0.094	R	0.082	S	0.066	T		1
17	234	S	0.766	T	0.204	A	0.009	F	0.009	P	0.004	R	0.004	S		
18	244	L	0.812	V	0.155	M	0.008	A	0.004	E	0.004	F	0.004	V		
19	242	R	0.545	K	0.240	S	0.161	T	0.037	A	0.012	Q	0.004	K		
20	246	L	0.736	V	0.191	I	0.061	E	0.004	R	0.004	X	0.004	L		
21	218	S	0.729	T	0.234	G	0.009	I	0.009	A	0.005	D	0.005	S		
22	217	C	0.991	R	0.005	S	0.005							C		
23	231	A	0.558	K	0.203	T	0.117	E	0.048	V	0.022	I	0.013	T		
24	235	A	0.638	V	0.174	G	0.064	I	0.055	T	0.030	F	0.026	A		
25	226	S	0.951	Y	0.027	F	0.009	C	0.004	K	0.004	T	0.004	S		
26	225	G	0.956	E	0.013	A	0.009	D	0.009	S	0.009	V	0.004	G		
27	213	F	0.559	Y	0.164	G	0.150	D	0.080	S	0.019	L	0.014	F		
28	203	T	0.571	S	0.286	I	0.049	N	0.049	P	0.015	A	0.005	N		1
29	207	F	0.749	V	0.111	I	0.068	L	0.053	T	0.010	A	0.005	I		1
30	202	S	0.762	T	0.119	N	0.035	G	0.020	R	0.020	A	0.010	K		1
31	199	S	0.482	T	0.136	D	0.104	N	0.087	G	0.060	K	0.040	D	H1	
32	202	Y	0.535	S	0.144	N	0.083	A	0.069	D	0.031	G	0.030	S	H1	
33	197	A	0.269	Y	0.162	G	0.147	W	0.117	S	0.091	T	0.066	Y	H1	
34	200	M	0.520	I	0.210	W	0.070	A	0.055	Y	0.050	V	0.040	M	H1	
35	196	S	0.372	H	0.235	N	0.077	A	0.061	G	0.051	Y	0.046	H	H1	
35a	33	-	0.824	W	0.096	V	0.043	G	0.016	S	0.016	N	0.005		H2	
35b	27	-	0.856	N	0.064	G	0.037	S	0.032	A	0.005	R	0.005		H3	
36	192	W	0.990	M	0.005	T	0.005							W		
37	193	V	0.741	I	0.228	L	0.021	G	0.005	Q	0.005			L		1
38	190	R	0.989	P	0.005	V	0.005							R		
39	190	Q	0.979	T	0.011	G	0.005	R	0.005					Q		
40	191	A	0.634	P	0.199	S	0.073	M	0.052	G	0.010	V	0.010	G		1
41	187	P	0.914	S	0.043	T	0.021	A	0.005	L	0.005	Q	0.005	P		
42	187	G	0.925	S	0.064	P	0.005	R	0.005					E		1
43	186	K	0.683	Q	0.183	R	0.124	E	0.005	H	0.005			Q		
44	186	G	0.882	A	0.048	S	0.043	R	0.027					G		
45	186	L	0.978	P	0.022									L		
46	185	E	0.956	Q	0.039	V	0.005							E		
47	184	W	0.989	S	0.011									W		

SUBSTITUTE SHEET (RULE 26)

14 / 28

48	185	V	0.481	M	2	I	0.173	L	0.124					I		
49	185	G	0.600	S	0.216	A	0.162	E	0.005	L	0.005	I	0.005	G		
50	185	R	0.146	W	0.146	V	0.119	A	0.114	G	0.081	Y	0.081	W	H2	
51	185	I	0.822	T	0.081	R	0.027	V	0.022	K	0.016	M	0.011	I	H2	
52	184	S	0.250	Y	0.239	N	0.123	K	0.060	I	0.054	D	0.050	D	H2	
52a	141	-	0.230	P	0.180	Y	0.153	G	0.126	N	0.066	V	0.055	P	H2	
52b	34	-	0.814	K	0.115	R	0.060	G	0.005	Y	0.005				H2	
52c	22	-	0.880	T	0.044	V	0.033	S	0.022	A	0.011	G	0.005		H2	
53	184	S	0.228	D	0.163	Y	0.125	G	0.109	N	0.082	H	0.054	E	H2	
54	183	G	0.328	S	0.202	D	0.129	N	0.112	K	0.082	F	0.055	N	H2	
55	182	G	0.544	S	0.181	D	0.085	W	0.066	Y	0.060	N	0.020	G	H2	
56	182	S	0.231	D	0.182	N	0.147	T	0.143	Y	0.077	G	0.060	D	H2	
57	184	T	0.582	K	0.120	N	0.065	A	0.054	I	0.054	P	0.022	T	H2	
58	183	Y	0.322	N	0.216	D	0.139	R	0.060	H	0.055	T	0.038	E	H2	
59	184	Y	0.908	F	0.043	N	0.016	S	0.011	D	0.005	G	0.005	Y	H2	
60	183	A	0.579	N	0.153	S	0.104	T	0.055	R	0.044	G	0.027	A	H2	
61	184	D	0.277	P	0.239	Q	0.174	A	0.141	V	0.076	T	0.033	P	H2	
62	185	S	0.686	K	0.146	P	0.065	N	0.038	G	0.016	R	0.016	K	H2	
63	186	V	0.511	L	0.247	F	0.215	S	0.011	A	0.005	K	0.005	F	H2	
64	186	K	0.581	Q	0.274	R	0.054	N	0.032	E	0.022	T	0.022	Q	H2	
65	186	G	0.688	S	0.237	T	0.032	A	0.016	D	0.011	E	0.011	G	H2	
66	186	R	0.935	Q	0.054	H	0.005	I	0.005					K		1
67	186	F	0.462	V	0.409	I	0.065	L	0.054	A	0.005	S	0.005	A		1
68	186	T	0.914	I	0.038	A	0.016	S	0.011	K	0.005	N	0.005	T		
69	187	I	0.791	M	0.139	V	0.032	D	0.005	F	0.005	G	0.005	F		1
70	187	S	0.684	T	0.214	N	0.070	L	0.032					T		
71	187	R	0.529	V	0.160	A	0.107	P	0.064	T	0.053	K	0.043	T		1
72	186	D	0.902	N	0.071	K	0.016	E	0.011					D		
73	185	T	0.368	N	0.266	D	0.177	K	0.070	E	0.059	A	0.011	T		
74	186	S	0.946	A	0.048	L	0.005							S		
75	187	K	0.674	T	0.139	I	0.070	R	0.027	A	0.021	F	0.021	S		1
76	187	N	0.701	S	0.251	K	0.027	R	0.011	T	0.005	Y	0.005	N		
77	187	T	0.615	Q	0.273	S	0.048	M	0.021	L	0.016	P	0.011	T		
78	186	L	0.364	A	0.273	F	0.235	V	0.096	I	0.005	M	0.005	A		
79	187	Y	0.638	S	0.239	F	0.059	V	0.048	H	0.005	M	0.005	Y		
80	187	L	0.782	M	0.207	N	0.005	-	0.005					L		
81	187	Q	0.529	E	0.205	K	0.122	R	0.032	T	0.032	N	0.027	Q		
82	194	M	0.497	L	0.421	W	0.051	V	0.015	I	0.010	-	0.005	L		
82a	195	N	0.442	S	0.291	R	0.077	T	0.066	D	0.053	G	0.020	S		
82b	194	S	0.795	N	0.082	R	0.051	G	0.026	T	0.021	A	0.010	S		
82c	197	L	0.701	V	0.234	M	0.041	G	0.010	A	0.005	D	0.005	L		
83	197	R	0.528	T	0.239	K	0.122	D	0.041	E	0.020	Q	0.015	T		
84	198	A	0.495	P	0.182	S	0.177	T	0.051	I	0.035	V	0.030	S		
85	198	E	0.591	A	0.172	D	0.126	S	0.051	V	0.045	G	0.015	E		
86	198	D	0.975	T	0.010	V	0.010	N	0.005					D		
87	198	T	0.929	S	0.035	G	0.010	M	0.010	A	0.005	Q	0.005	T		
88	198	A	0.939	G	0.040	P	0.005	T	0.005	V	0.005	Y	0.005	A		
89	198	V	0.768	L	0.066	M	0.056	T	0.045	I	0.040	F	0.010	V		
90	199	Y	0.980	F	0.010	A	0.005	I	0.005					Y		
91	199	Y	0.930	F	0.045	C	0.015	R	0.005	T	0.005			Y		
92	198	C	0.990	A	0.005	M	0.005							C		
93	198	A	0.838	T	0.076	V	0.061	H	0.005	K	0.005	N	0.005	N		1
94	198	R	0.596	K	0.162	T	0.051	G	0.045	P	0.045	Q	0.025	E		1
95	161	G	0.174	D	0.120	E	0.099	A	0.093	N	0.092	P	0.068	G		
96	159	P	0.168	R	0.130	G	0.112	L	0.062	V	0.062	Y	0.062	T	H3	
97	156	G	0.170	P	0.094	V	0.094	E	0.088	T	0.069	S	0.063	P	H3	
98	155	G	0.152	Y	0.101	L	0.095	D	0.087	V	0.076	S	0.063	T	H3	

SUBST ~~FIG. 2A-2~~ (RULE 26)



15 / 28

99	143	G	0.172	Y	0.108	T	0.102	-	0.089	A	0.076	E	0.070	G	H3	
100	131	-	0.171	S	0.165	Y	0.146	G	0.095	V	0.070	R	0.051	P	H3	
100a	110	-	0.304	G	0.146	S	0.095	D	0.046	A	0.044	L	0.044	Y	H3	
100b	99	-	0.369	G	0.134	S	0.127	T	0.076	Y	0.045	V	0.038	Y	H3	
100c	92	-	0.410	G	0.122	Y	0.103	D	0.058	S	0.058	P	0.045		H3	
100d	72	-	0.538	Y	0.058	G	0.051	S	0.051	C	0.045	L	0.038		H3	
100e	62	-	0.600	Y	0.155	S	0.045	F	0.032	G	0.032	A	0.026		H3	
100f	53	-	0.658	Y	0.097	H	0.039	R	0.039	P	0.026	S	0.026		H3	
100g	41	-	0.735	Y	0.084	G	0.065	Q	0.026	S	0.019	D	0.013		H3	
100h	30	-	0.806	Y	0.058	D	0.032	A	0.019	G	0.019	S	0.019		H3	
100i	24	-	0.844	Y	0.039	G	0.026	X	0.019	L	0.013	N	0.013		H3	
100j	80	-	0.481	Y	0.149	A	0.117	W	0.084	F	0.045	G	0.039		H3	
100k	138	F	0.503	M	0.144	L	0.137	-	0.098	D	0.039	V	0.033	F	H3	
101	149	D	0.754	A	0.073	R	0.066	N	0.020	Q	0.020	P	0.013	D	H3	
102	151	Y	0.368	V	0.224	I	0.112	S	0.086	P	0.072	H	0.053	Y	H3	
103	154	W	0.955	E	0.013	F	0.013	D	0.006	R	0.006	Y	0.006	W		
104	154	G	0.974	Y	0.013	D	0.006	T	0.006					G		
105	154	Q	0.798	R	0.104	K	0.045	E	0.013	N	0.013	S	0.013	Q		
106	155	G	0.987	Y	0.006	-	0.006							G		
107	152	T	0.908	S	0.026	V	0.020	G	0.013	I	0.007	L	0.007	T		
108	152	L	0.645	T	0.178	M	0.105	P	0.020	K	0.013	R	0.013	T		
109	151	V	0.967	L	0.013	I	0.007	M	0.007	X	0.007			V		
110	151	T	0.940	S	0.026	I	0.013	A	0.007	H	0.007	V	0.007	T		
111	137	V	0.978	I	0.015	T	0.007							V		
112	138	S	0.971	T	0.014	R	0.007	V	0.007					S		
113	131	S	0.962	P	0.015	A	0.008	L	0.008	T	0.008			S		

**FIG.\_ 12A-3**

16 / 28

Pos. Light Chain	Number of Observations	Observed Frequencies of 5 Most Abundant Amino Acids in Alignment of Human Sequences												CAB1 Sequence	CDR	Mutated Residues
1	95	Q	0.589	S	0.158	N	0.095	H	0.074	D	0.053	F	0.021	E		1
2	139	S	0.446	Y	0.388	F	0.101	V	0.043	L	0.014	T	0.007	N		1
3	140	V	0.307	E	0.243	A	0.207	M	0.093	D	0.064	I	0.043	V		
4	140	L	0.971	V	0.029									L		
5	141	T	0.915	A	0.021	S	0.021	I	0.014	K	0.007	L	0.007	T		
6	140	Q	0.993	E	0.007									Q		
7	139	P	0.906	D	0.029	S	0.029	A	0.022	E	0.014			S		1
8	139	P	0.741	A	0.137	H	0.072	R	0.029	L	0.007	S	0.007	P		
9	139	S	0.964	A	0.014	V	0.014	R	0.007					A		1
10	0	-	1.000											I		1
11	138	V	0.790	A	0.138	L	0.058	M	0.014					M		1
12	139	S	0.978	F	0.007	T	0.007	E	0.004	Q	0.004			S		
13	138	V	0.406	G	0.348	A	0.138	E	0.087	L	0.014	D	0.007	A		
14	135	S	0.630	A	0.230	T	0.111	D	0.007	F	0.007	G	0.007	S		
15	135	P	0.881	L	0.089	A	0.022	S	0.007					P		
16	134	G	0.978	E	0.015	L	0.007							G		
17	133	Q	0.811	K	0.098	A	0.045	E	0.024	G	0.015	H	0.008	E		1
18	133	T	0.504	S	0.263	R	0.135	K	0.068	E	0.008	G	0.008	K		1
19	130	V	0.454	A	0.385	I	0.146	G	0.008	L	0.008			V		
20	128	T	0.531	R	0.188	S	0.148	K	0.047	I	0.031	M	0.016	T		
21	121	I	0.901	V	0.050	L	0.017	A	0.008	F	0.008	M	0.008	I		
22	120	S	0.492	T	0.475	A	0.008	G	0.008	I	0.008	N	0.008	T		
23	117	C	1.000											C		
24	112	S	0.536	T	0.259	G	0.089	A	0.045	Q	0.033	I	0.018	S	L1	
25	108	G	0.870	L	0.056	R	0.028	A	0.019	I	0.009	P	0.009	A	L1	
26	108	D	0.339	S	0.250	T	0.213	N	0.087	E	0.037	G	0.037	S	L1	
27	104	S	0.415	N	0.118	K	0.113	A	0.104	T	0.066	G	0.047	S	L1	
28	104	L	0.346	S	0.346	I	0.115	G	0.067	A	0.058	D	0.019	S	L1	
29	100	G	0.243	N	0.239	D	0.159	S	0.078	P	0.068	H	0.058	V	L1	
30	103	I	0.291	V	0.165	D	0.136	N	0.107	E	0.058	S	0.049	S	L1	
31	101	G	0.356	K	0.168	A	0.099	E	0.084	Q	0.084	D	0.069	Y	L1	
31a	54	-	0.438	S	0.167	G	0.104	N	0.083	Y	0.063	D	0.052	M	L1	
31b	49	-	0.495	N	0.227	Y	0.155	S	0.041	G	0.021	H	0.021	H	L1	
31c	23	-	0.760	N	0.134	S	0.031	K	0.021	D	0.012	E	0.010		L1	
31d	0	-	1.000												L1	
31e	0	-	1.000												L1	
31f	0	-	1.000												L1	
32	94	Y	0.515	S	0.134	F	0.093	A	0.072	T	0.052	H	0.041		L1	
33	97	V	0.680	A	0.186	I	0.082	Y	0.021	F	0.010	P	0.010		L1	
34	92	S	0.380	H	0.120	A	0.109	Y	0.098	N	0.076	Q	0.076		L1	
35	98	W	0.990	Y	0.010									W		
36	96	Y	0.844	F	0.073	H	0.073	W	0.010					F		1
37	95	Q	0.916	R	0.042	E	0.011	H	0.011	K	0.011	Y	0.011	Q		
38	94	Q	0.862	H	0.053	L	0.053	E	0.011	K	0.011	V	0.011	Q		
39	93	K	0.333	L	0.172	R	0.161	H	0.151	Q	0.086	V	0.043	K		
40	93	P	0.946	S	0.022	A	0.011	L	0.011	R	0.011			P		

**FIG. 12B-1**  
SUBSTITUTE SHEET (RULE 26)

17 / 28

41	93	G	0.871	H	0.000	D	0.022	R	0.022	P	0.0	V	0.011	G		
42	92	Q	0.424	T	0.217	K	0.163	R	0.087	S	0.054	G	0.022	T		
43	92	A	0.717	S	0.174	G	0.065	T	0.022	L	0.011	V	0.011	S		
44	93	P	0.978	A	0.011	M	0.011							P		
45	92	K	0.391	V	0.315	R	0.109	L	0.065	T	0.065	A	0.033	K		
46	92	L	0.728	V	0.076	F	0.065	T	0.043	A	0.022	M	0.022	L		
47	91	V	0.484	L	0.374	I	0.077	M	0.055	N	0.011			W		1
48	91	I	0.791	V	0.110	M	0.077	L	0.011	S	0.011			I		
49	91	Y	0.769	F	0.110	R	0.066	H	0.022	D	0.011	I	0.011	Y		
50	89	D	0.303	E	0.210	Q	0.093	V	0.067	G	0.056	K	0.056	S	L2	
51	88	D	0.364	N	0.205	V	0.159	H	0.068	T	0.068	G	0.034	T	L2	
52	89	N	0.393	T	0.213	S	0.202	D	0.101	A	0.022	F	0.011	S	L2	
53	88	K	0.307	D	0.193	Q	0.182	N	0.080	E	0.057	S	0.057	N	L2	
54	88	R	0.875	X	0.068	K	0.034	L	0.011	W	0.011			L	L2	
55	86	P	0.851	G	0.080	S	0.023	A	0.011	H	0.011	R	0.011	A	L2	
56	85	S	0.837	D	0.081	P	0.023	A	0.012	L	0.012	T	0.012	S	L2	
57	86	G	0.920	E	0.034	S	0.011	T	0.011	W	0.011	-	0.011	G		
58	84	I	0.600	V	0.353	A	0.012	G	0.012	T	0.012	-	0.012	V		
59	84	P	0.847	S	0.106	A	0.012	L	0.012	V	0.012	-	0.012	P		
60	85	D	0.488	E	0.325	N	0.047	A	0.035	H	0.023	L	0.023	A		1
61	87	R	0.977	D	0.011	-	0.011							R		
62	88	F	0.943	I	0.034	L	0.011	R	0.011					F		
63	87	S	0.989	F	0.011									S		
64	87	G	0.885	A	0.069	S	0.023	V	0.023					G		
65	87	S	0.977	G	0.011	Y	0.011							S		
66	86	K	0.430	N	0.186	S	0.186	T	0.081	X	0.070	R	0.035	G		1
67	85	S	0.953	T	0.024	K	0.012	L	0.012					S		
68	85	G	0.859	S	0.071	A	0.035	D	0.024	Q	0.012			G		
69	85	N	0.434	T	0.318	A	0.129	D	0.036	G	0.024	K	0.024	T		
70	85	T	0.529	S	0.341	E	0.082	A	0.024	K	0.024			S		
71	85	A	0.847	R	0.082	V	0.059	S	0.012					Y		1
72	85	T	0.447	S	0.424	Y	0.082	A	0.035	I	0.012			S		
73	85	L	0.988	S	0.012									L		
74	85	T	0.706	A	0.165	G	0.106	I	0.012	L	0.012			T		
75	85	I	0.929	V	0.047	A	0.012	L	0.012					I		
76	85	S	0.718	T	0.200	N	0.035	I	0.024	G	0.012	R	0.012	S		
77	85	G	0.765	R	0.129	S	0.094	E	0.012					R		
78	85	L	0.588	V	0.224	T	0.106	A	0.071	G	0.012			M		1
79	85	Q	0.659	E	0.153	R	0.071	K	0.047	L	0.024	A	0.012	E		
80	85	A	0.459	S	0.235	T	0.200	V	0.047	P	0.035	N	0.012	A		
81	85	E	0.541	G	0.235	M	0.071	D	0.047	L	0.024	N	0.024	E		
82	85	D	0.964	N	0.024	E	0.012							D		
83	85	E	0.976	D	0.012	T	0.012							A		1
84	85	A	0.941	T	0.035	E	0.012	S	0.012					A		
85	85	D	0.859	E	0.082	H	0.024	A	0.012	I	0.012	M	0.012	T		1
86	85	Y	0.976	F	0.012	H	0.012							Y		
87	85	Y	0.894	F	0.106									Y		
88	85	C	0.988	H	0.012									C		
89	85	Q	0.482	A	0.153	S	0.141	G	0.094	C	0.059	N	0.035	Q	L3	
90	85	S	0.388	T	0.271	A	0.212	V	0.118	L	0.012			Q	L3	
91	85	W	0.576	Y	0.247	A	0.059	F	0.035	R	0.035	D	0.012	R	L3	
92	84	D	0.606	G	0.095	A	0.071	N	0.061	T	0.048	E	0.024	S	L3	
93	84	S	0.405	D	0.179	G	0.107	N	0.095	P	0.071	T	0.060	S	L3	
94	84	S	0.536	G	0.155	N	0.073	R	0.060	D	0.058	T	0.048	Y	L3	
95	82	S	0.265	L	0.253	G	0.108	N	0.096	T	0.084	A	0.036	P	L3	

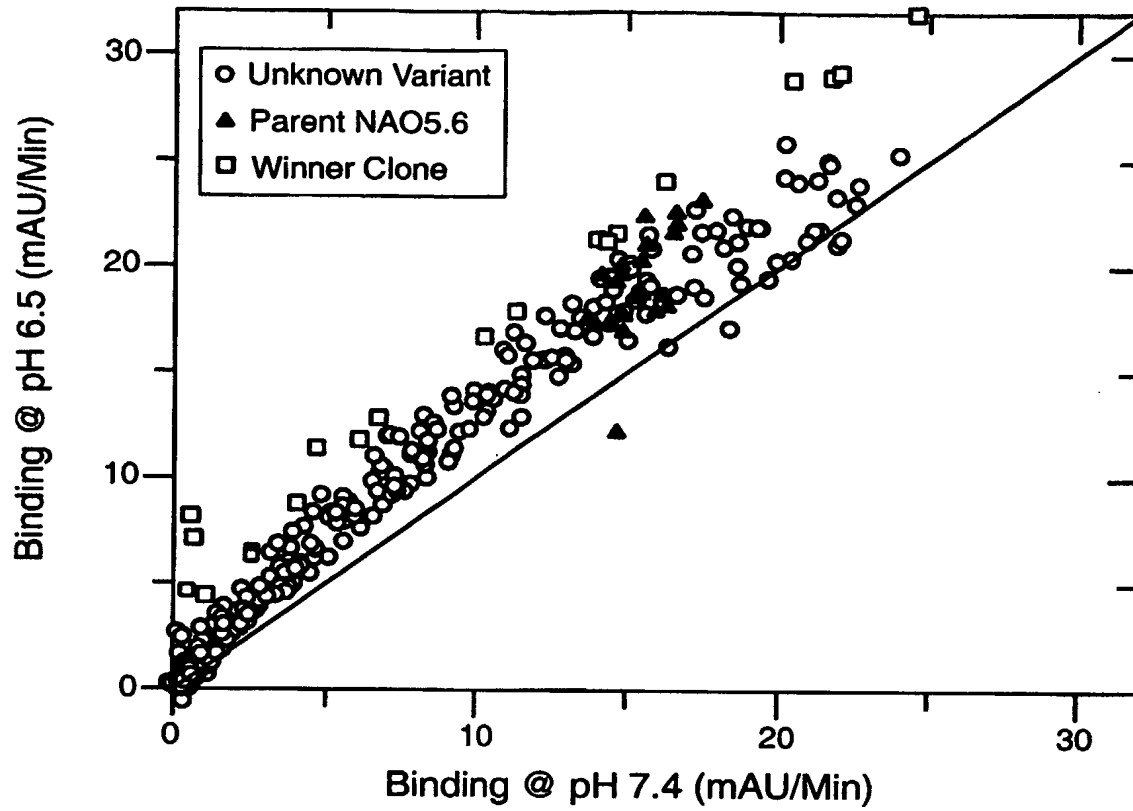
SUBFIG 12B-2 (RULE 26)

18 / 28

95a	60	-	0.268	S	0.183	D	0.159	N	0.110	T	0.073	Q	0.049	L	L3	
95b	40	-	0.512	A	0.098	G	0.098	H	0.085	E	0.049	R	0.037	T	L3	
95c	5	-	0.939	P	0.037	A	0.012	G	0.012						L3	
95d	1	-	0.988	G	0.012										L3	
95e	0	-	1.000												L3	
95f	0	-	1.000												L3	
96	80	V	0.305	G	0.098	P	0.098	W	0.098	A	0.073	N	0.073		L3	
97	85	V	0.788	I	0.118	L	0.047	M	0.035	G	0.012				L3	
98	86	F	0.988	V	0.012									F		
99	89	G	0.989	F	0.011									G		
100	89	G	0.831	T	0.124	A	0.022	S	0.022					A		1
101	89	G	1.000											G		
102	89	T	0.989	G	0.011									T		
103	88	K	0.739	N	0.091	R	0.068	Q	0.034	T	0.034	E	0.011	K		
104	87	L	0.667	V	0.322	Q	0.011							L		
105	87	T	0.954	S	0.023	I	0.011	L	0.011					E		1
106	85	V	0.988	T	0.012									L		1
106a	84	L	0.952	V	0.024	P	0.012	Q	0.012					K		1
107	78	G	0.782	S	0.103	R	0.090	C	0.013	L	0.013			R		1
108	46	Q	0.957	P	0.022	R	0.022							A		1
109	46	P	0.957	K	0.022	Q	0.022							A		1

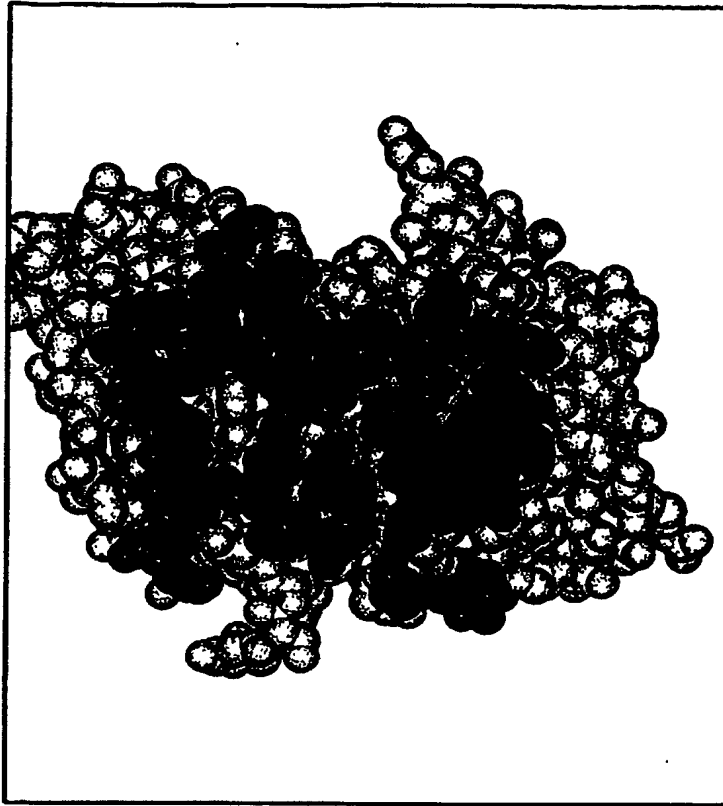
**FIG.\_ 12B-3**

19 / 28

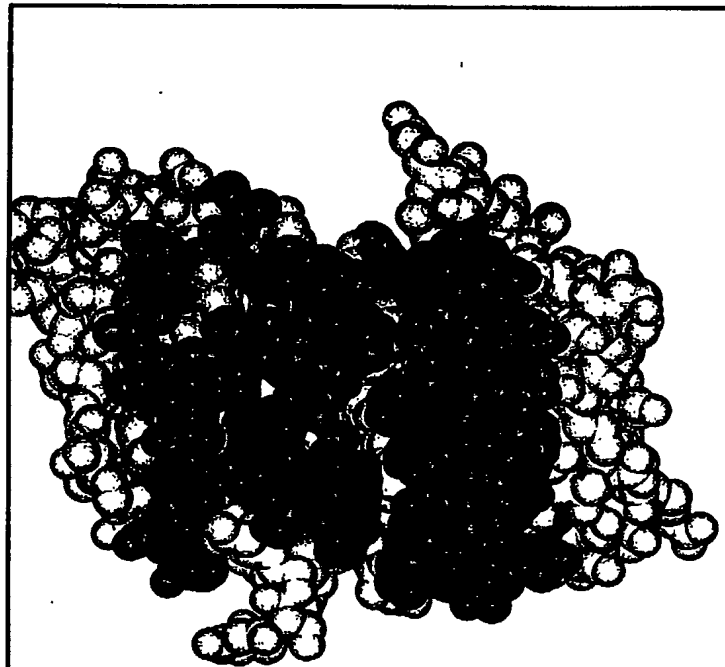
**FIG.\_13**

20 / 28

Residues Chosen for  
Mutagenesis in Na06.6



CDRs in Na06.6



**FIG. 14**

SUBSTITUTE SHEET (RULE 26)

21 / 28

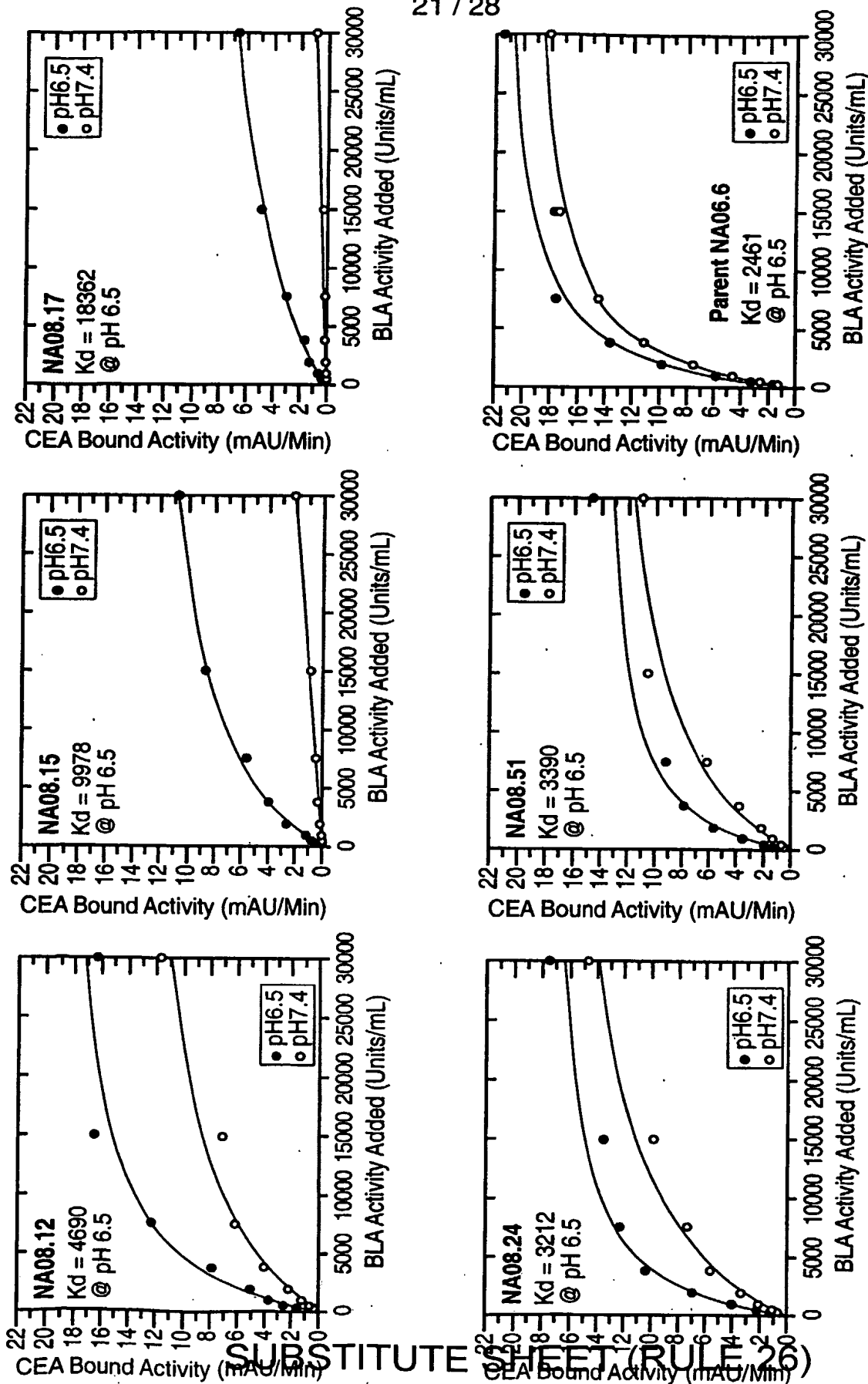


FIG. 1E

SUBSTITUTE SHEET (RULE 26)

22 / 28

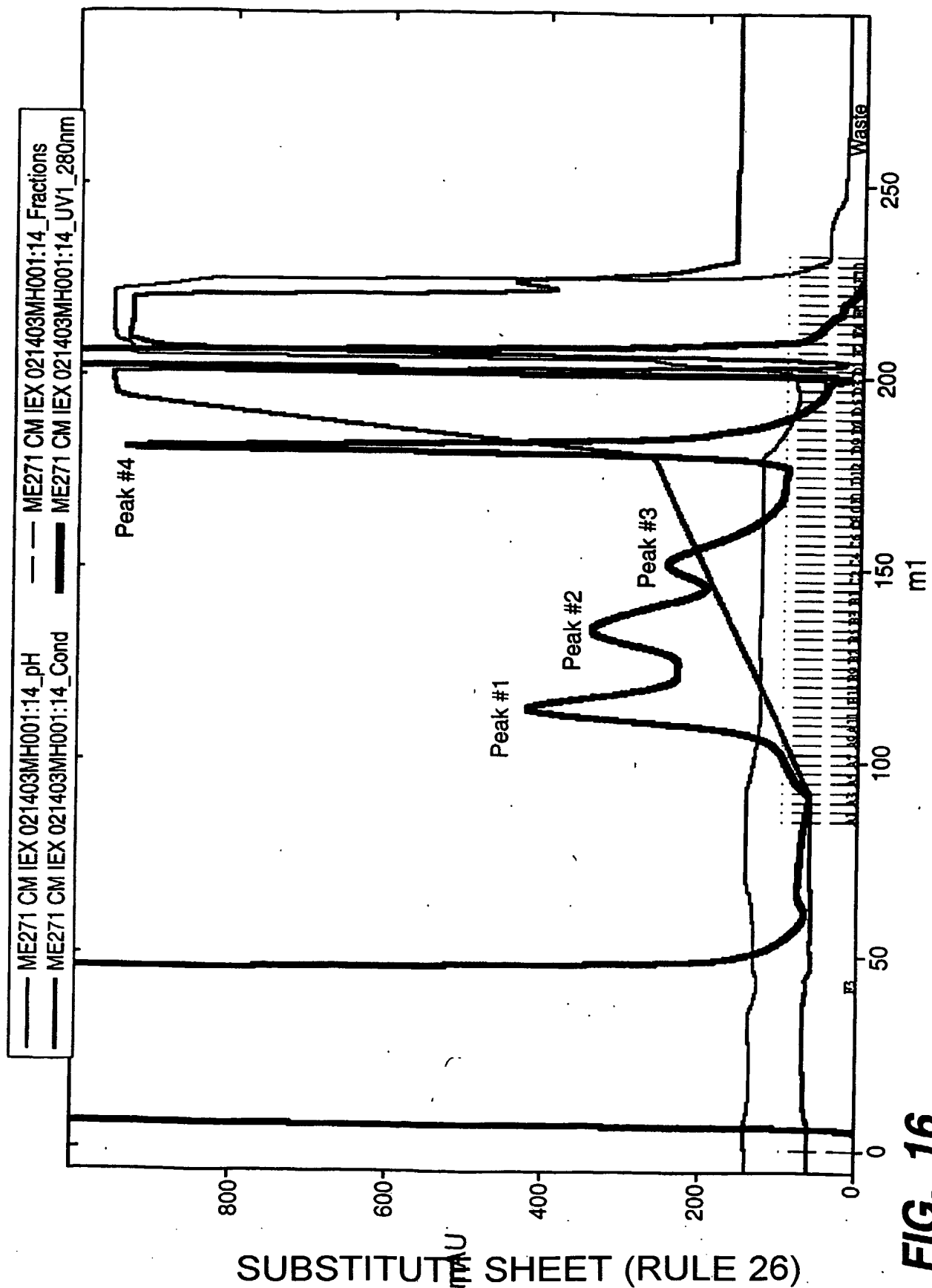
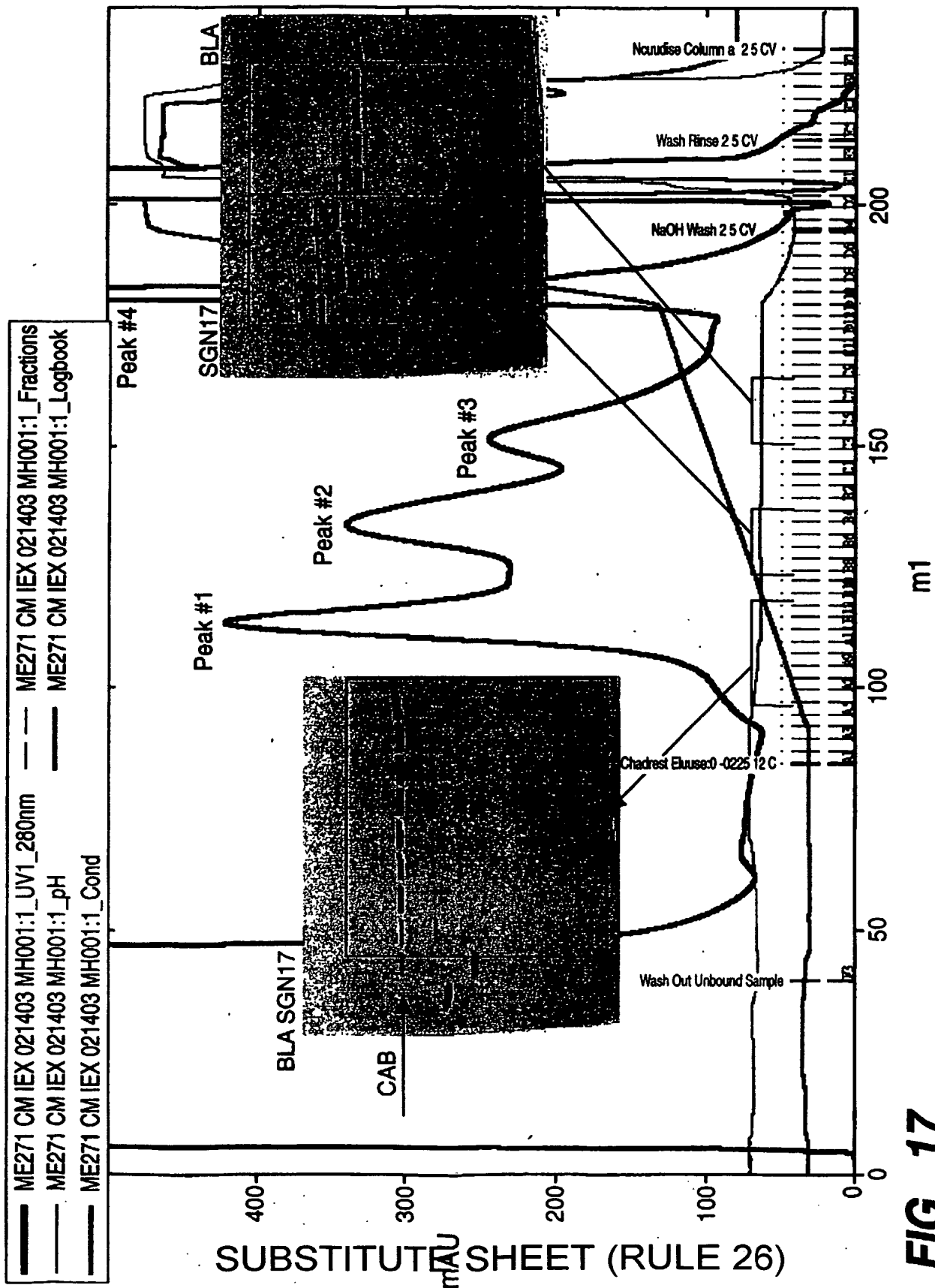


FIG. 16





24 / 28

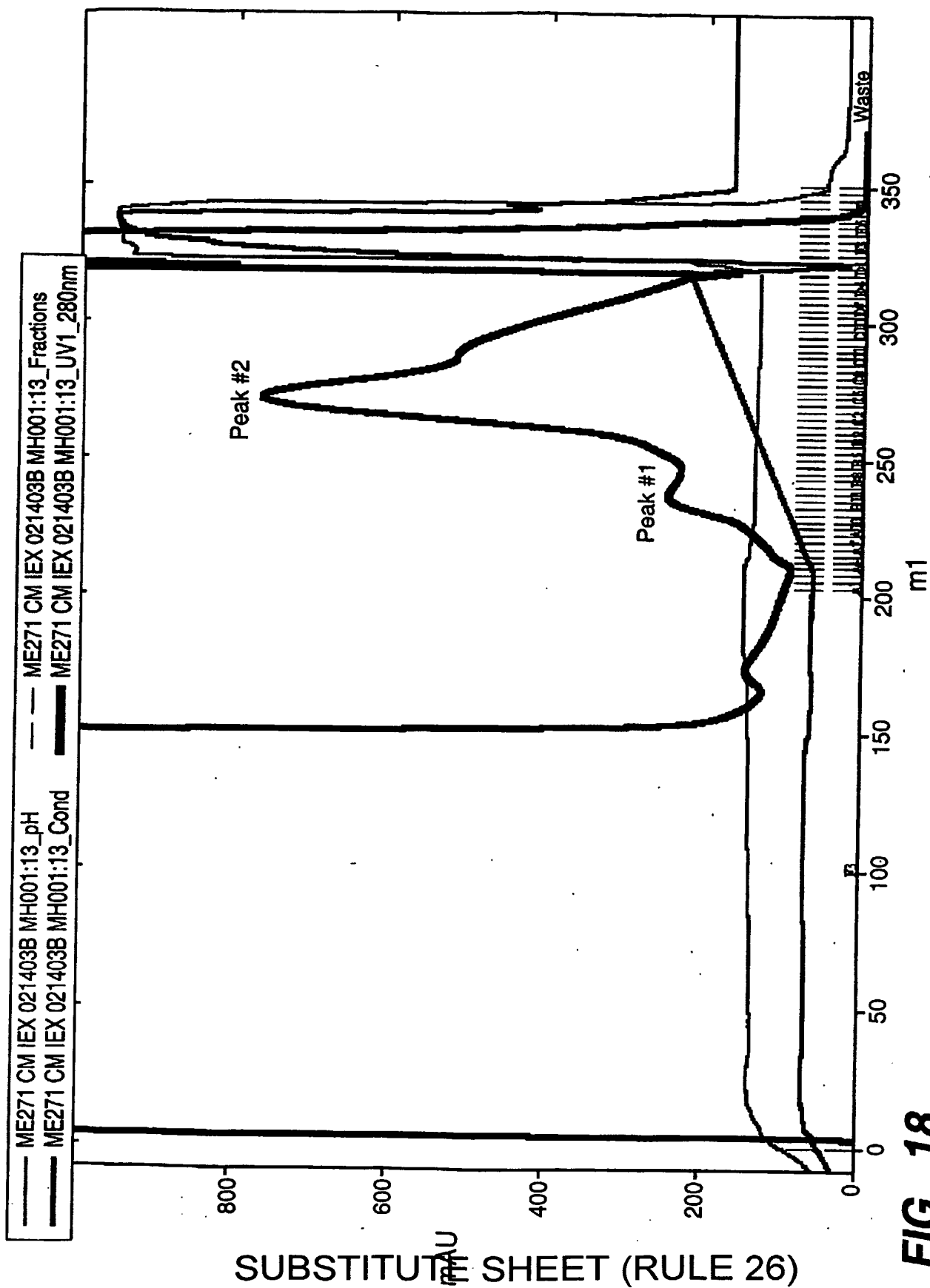


FIG.-18

25 / 28

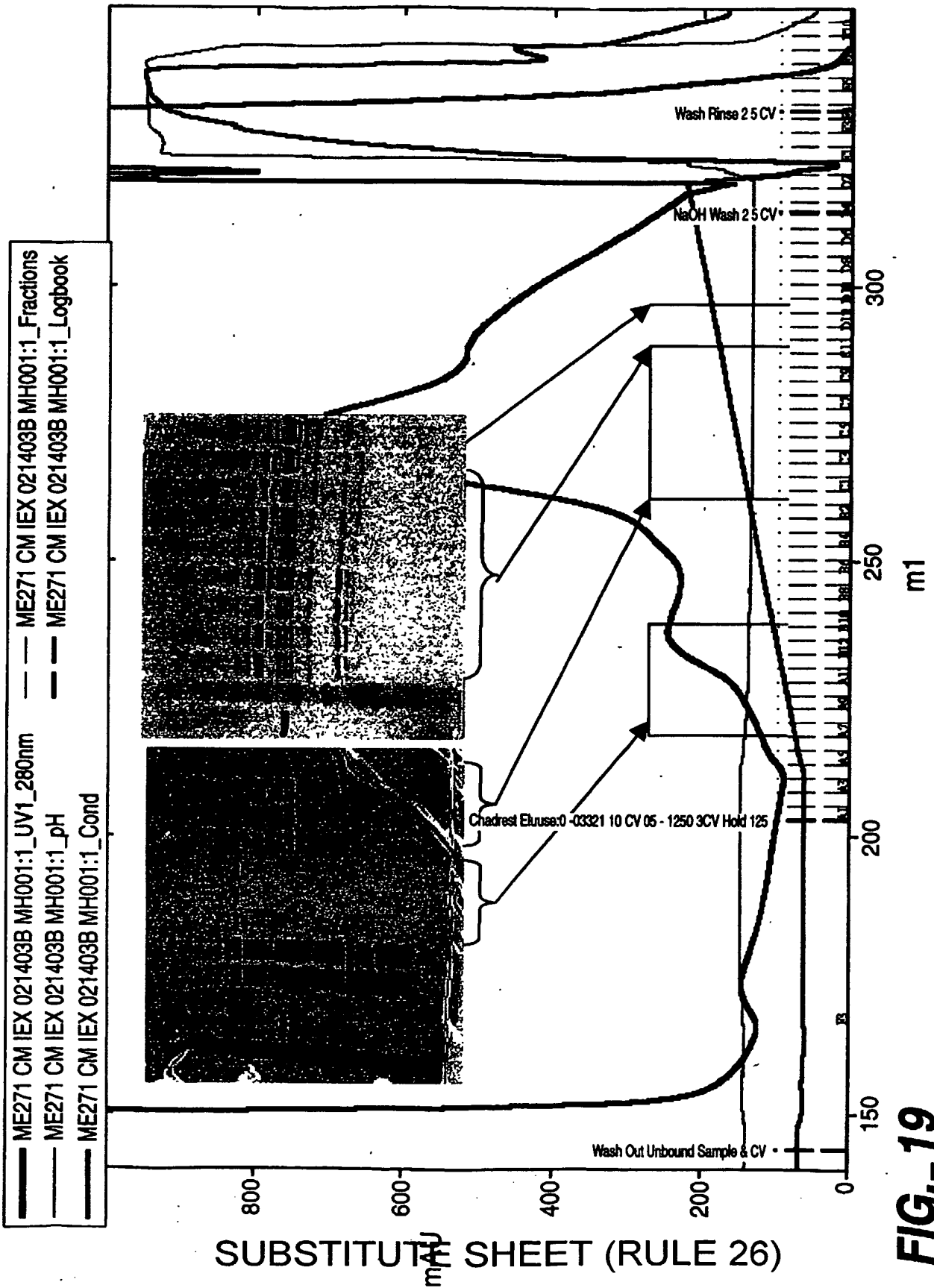


FIG. 19

SUBSTITUTE SHEET (RULE 26)

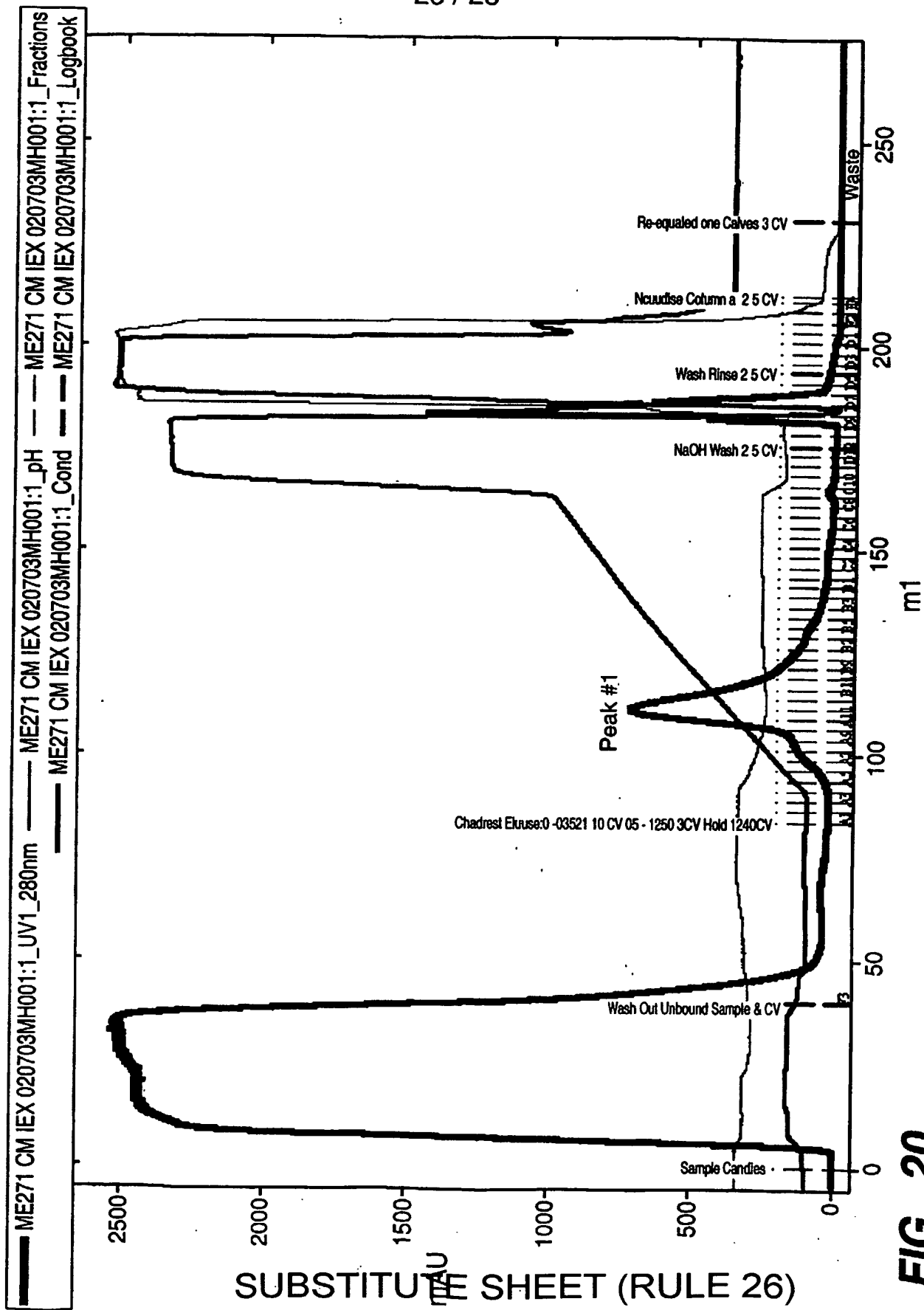
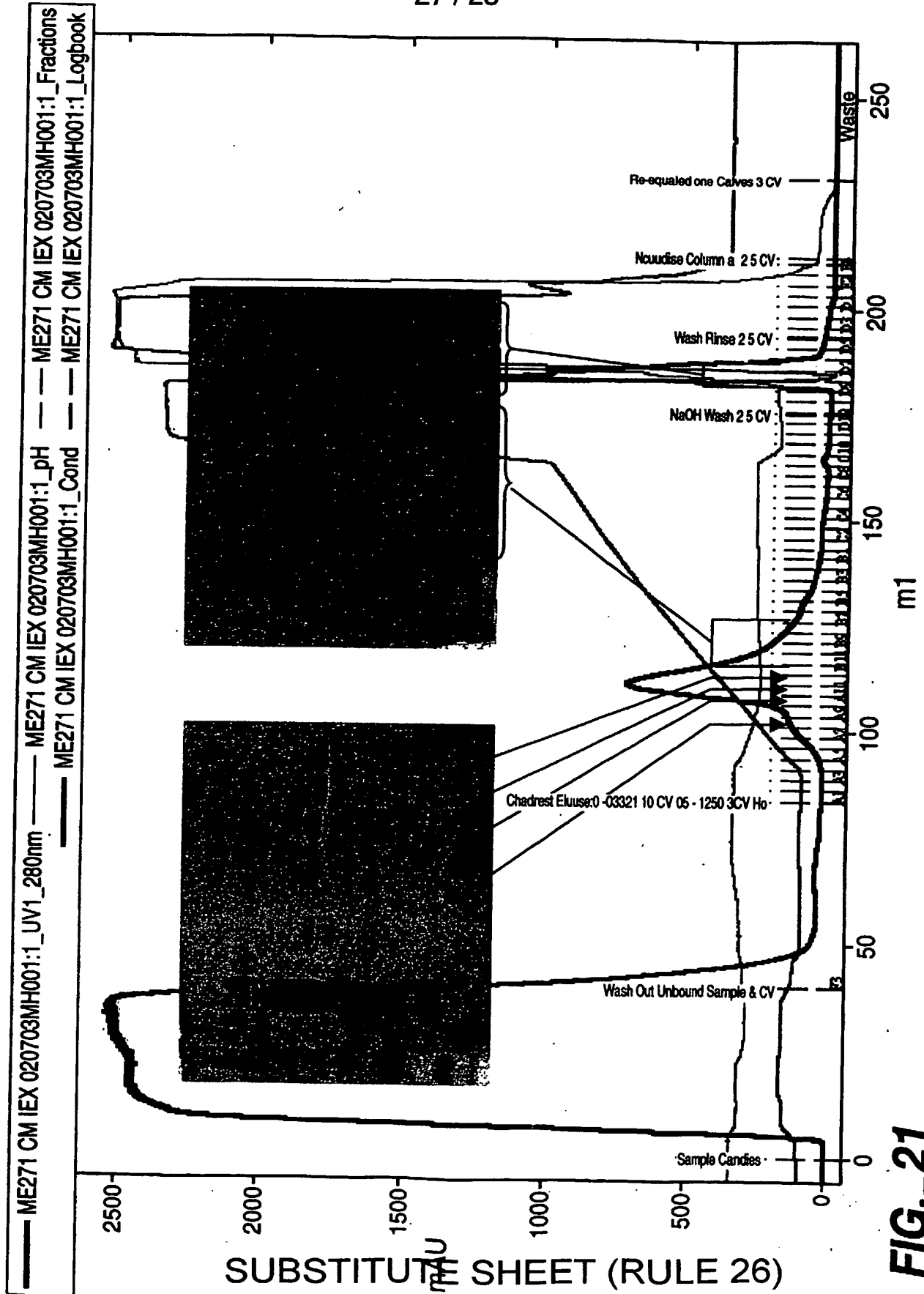


FIG. 20



SUBSTITUTE SHEET (RULE 26)

FIG. 21

28 / 28

**PBA Purification**

Feed FT Wash Eluate



***FIG.\_22***